Author:
Devitt, Laura

Title:
Evaluation of a hydrological modelling framework, DECIPHeR, for use in large and data scarce river basins

Upper Niger case study

EVALUATION OF A HYDROLOGICAL MODELLING FRAMEWORK, DECIPHeR, FOR USE IN LARGE AND DATA SCARCE RIVER BASINS: UPPER NIGER CASE STUDY

LAURA DEVITT

OCTOBER 2018

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Geography (MSc) (R) in the Faculty of Science, School of Geographical Sciences, October 2018

WORD COUNT = 27095

ABSTRACT

Hydrological models are important tools for making predictions of river flows to be used in water resources management strategies, and for hypotheses testing in hydrological research. Hydrological modelling studies have been constrained by data availability and computational efficiency in the past; however, the development of national and global open-source data products and significant gains in computational power has allowed hydrological models to be implemented on many spatial scales. DECIPHeR (Dynamic fluxEs and ConnectIvity for Predictions of Hydrology) is a new flexible hydrological modelling framework that simulates streamflow from spatial scales of small headwaters catchments to entire continents. DECIPHeR can be adapted by the user to specific hydrological systems and data availability, and can be modified to represent different hydrological processes. Here, DECIPHeR is applied to the Upper Niger River in West Africa – a large and data scarce basin. This basin is characterised by highly variable climatic and physiographic features from its upstream headwaters to the downstream outlet. The initial DECIPHeR model structure was applied across the Upper Niger basin, forced with one global precipitation data product (MSWEP), and three global PET data products (GLEAM, ECMWF EartH2Observe, and temperature-based estimates). Model performance was evaluated with three performance metrics (NSE, PBIAS, and low-flow volume bias). Initial simulations were able to reproduce the shape and timing of the flow peaks, but were overpredicting flow volumes by a factor of four in all seven sub-basins. The model structure was modified to include a module to represent evaporation from the saturated zone and applied across the basin. This improved model performance significantly in all sub-basins. However, 'behavioural' (i.e. NSE>0.5, PBIAS<10, low-flow volume bias<10) model simulations could only be identified in three downstream sub-basins. This is due to large water balance issues in the headwater catchments, likely caused by large errors in the global input data products and a lack of information about the processes occurring in these sub-basins. However, this study has shown that DECIPHeR is particularly valuable for modelling studies in domains where there are little or no ground observations to inform our understanding of catchment functioning, and allows for multiple hypotheses of the dominant processes to be tested.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Professor Jim Freer, Dr Jeff Neal, and Dr Gemma Coxon for all of their support and invaluable advice throughout this year. This has been the most challenging project I have undertaken to date and would not have been possible without their guidance, help and positivity.

A special thank you should also be given to Gemma for teaching me how to use the model and helping me get to grips with it, and for all her technical and programming help.

I would also like to thank the University of Bristol School of Geographical Sciences for what has been the most enjoyable four years. Lastly, thank you to the Browns MScR girls for always being there for me and being so supportive throughout the whole process.

AUTHOR DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE:..........................

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The earth's water resources require management in order to meet the demands of present and future generations, and to reduce the consequences of hydrological extremes. Accurate simulations and predictions of flows in rivers are increasingly required at many temporal and spatial scales for many different management strategies, such as adaptation to climate change and mitigation of extreme events, such as floods and droughts (Singh and Woolhiser, 2002; Son and Sivapalan, 2007; Mockler et al., 2016). Hydrological models are a valuable tool for both water resources management and research, as they are used to help inform decisions, advance our knowledge of the hydrological system (Zhang et al., 2008), and to assess the impacts and uncertainty of changes (e.g. climate and land-use change) (Wheater et al., 1993; Dunn et al., 2008; Zhang et al., 2008).

In the past, hydrological modelling has been constrained by data availability, with studies mainly being carried out in the developed world at the catchment scale. However, the development of, and greater access to, many national and global open-source datasets, and significant gains in computing power, has allowed hydrological models to be implemented on many spatial scales, from national, to continental, and even globally (e.g. Gudmundsson et al., 2012; Kauffeldt et al., 2016; McMillan et al., 2016). Much of the developing world is particularly vulnerable to future changes of the climate, and has been predicted to see an intensification of the hydrological cycle (Diallo et al., 2016) and therefore being able to develop and implement hydrological models in these regions will aid in mitigation and adaptation strategies.

The increasing threats posed by future climate change and anthropogenic activities on the hydrological cycle has drawn attention of the modelling community to providing predictions of future dynamics and impacts on large spatial scales. Large-scale hydrology is defined as encompassing spatial scales that are greater than a single catchment all the way to the global-scale (Cloke and Hannah, 2011). Large-scale hydrological modelling studies are needed in order to determine the driving forces, patterns, and feedback responses of real world systems, and this information can be used to identify regions that are most vulnerable to changes and intensifications of the water cycle. However, there are inevitably obstacles and uncertainty when using large-scale models, e.g. choosing model structures and parameters that can describe a large number of catchments with varying regimes and properties.

There is a large variety of hydrological models available, with varying levels of spatial complexity and different representations of the dominant processes and catchment behaviour. However, there is little consensus amongst the hydrological research community on the appropriate model that should be used for different objectives. A 'perfect' model of the hydrological system does not exist, and therefore all models are in error (Freer et al., 2004). Traditionally, hydrological model developments have largely employed a 'one-size-fits-all' approach, aiming to build a single model structure to implement in all catchments (Fenicia et al., 2014). However, hydrological processes and responses to external drivers vary in catchments across the globe. Attempting to represent spatial variability and heterogeneity of catchment characteristics (McDonnell et al., 2007), the perception of ‘uniqueness of place’ (Beven, 2000), creating a model that ‘works for the right reasons’ when evaluated (Kirchner, 2006), and the impact of quality, resolution an availability of input and output data, with one model structure is an extremely complex task (Schoups et al., 2008; Kavetski et al., 2011).

1.1. DECIPHeR

DECIPHeR (Dynamic fluxEs and ConnectIvity for Predictions of HydRology, Coxon et al., 2018) is a new flexible hydrological modelling framework for uncertain flow simulation and prediction at catchment to continental scales. It builds on the key concepts of Dynamic TOPMODEL, originally developed by Beven and Freer (2001a). The model can be modified and adapted to suit specific hydrological settings and available data for the region of interest. During the model set-up, the user is given a number of choices in the way that they wish to represent different levels of spatial heterogeneity, connectivity and hydrological processes. These decisions are ultimately based on the data availability, the user’s understanding of the model domain of interest, and what is appropriate for the specific task. DECIPHeR has been applied in the UK as a demonstration for use in a large-scale application and initial evaluation of model performance (Coxon et al., 2018). In this first implementation, only one model structure was applied homogenously across the domain. In addition, the UK is a data rich and well gauged location. Although this initial application serves as a good benchmark of DECIPHeR’s ability at large scales, further studies are required that evaluate its performance and capacity for modelling in large and data scarce river basins.

1.2. THE NIGER BASIN

The Niger River is the largest river is West Africa (approximately 2.27 km^2), with more than 100 million inhabitants, many of whom are dependent on the river for their livelihoods. The river extends across nine countries, flowing through several distinct climatic zones, ranging from tropical humid in the Guinea Highlands, to the Sahara desert. In recent decades, there has been a drastic decrease in the precipitation in the Sahelian belt, and this has caused major disasters, such as droughts and famine (Mahe et al., 2013). Recently, flooding is becoming a growing concern in the river basin, and this has been attributed to climate change and changes in land use (Aich et al., 2015, 2016). Western Africa is also a region that is particularly vulnerable and sensitive to future climate change (Diallo et al., 2016), and therefore may begin to see more of an intensification of the hydrological cycle and more extreme events. Therefore, a better understanding of the hydrological responses in the basin could help to improve water resources management, infrastructure designs, and the development of robust operational flood and drought forecasts (Andersson et al., 2017a, b). Hydrological modelling in this basin is a valuable tool for these objectives. Previous model studies have proved to be successful in reproducing the annual and seasonal dynamics of the Niger's river flows, however, there is a tendency for simulations to overpredict the magnitude of these flows (e.g. Schoul and Abbaspour, 2006; Pedinotti et al., 2012; Andersson et al., 2017a). This indicates that the model structures used are not representing the dominant processes occurring in the basin.

1.3. SCOPE OF DISSERTATION AND RESEARCH QUESTIONS

DECIPHeR is applied to the Upper Niger basin in West Africa in this study. This catchment was chosen as a suitable location for investigating model performance in large and data scarce locations, and for highlighting areas for future model development because: 1) there are very few ground observations of catchment functioning to inform model structural choices, 2) the data that is used is from open-source global products, and this study provides an evaluation of these datasets in hydrological modelling at larger scales, and 3) there is a strong gradient in hydroclimatic variability from the upstream headwaters to the downstream outlet, which provides a test of the rainfall-runoff model's capacity to represent these different catchment dynamics. Therefore, this dissertation aims to:

1. Evaluate the performance of a new modelling framework, DECIPHeR, in a large and data scarce model domain, using the Upper Niger basin as a case study.
2. Investigate the different sources of uncertainty (input, parametric and model structure uncertainty) and analyse the impacts that these have on the hydrological model's performance.
3. Highlight areas of this new framework where future improvements and developments are required for modelling at large scales.

2. LITERATURE REVIEW

This section aims to review some key research themes that are relevant to the research aims and questions that frame this hydrological modelling study. Firstly, the large variety of available hydrological models that are available is discussed, with the advantages and disadvantages of each approach given. Within this, the increased interest of modelling at large scales is discussed, with the associated limitations and challenges. Next, the importance of choosing appropriate model evaluation techniques is highlighted, followed by a discussion of the key sources of uncertainty that are present in hydrological modelling. Finally, the challenges of modelling in large and data scarce river basins are introduced, and the key findings of past modelling studies conducted in the Upper Niger basin are summarised.

2.1. HYDROLOGICAL MODELLING

Hydrological models are an important tool to help inform decisions in all aspects of water resources management. Robust simulations and predictions of river discharges are needed for many management strategies, such as local-scale flood and drought mitigation and prevention, to the assessment of available freshwater to meet population demands at regional to continental scales (Archfield et al., 2015). There is a large variety of hydrological models available that have been built to address a wide range of issues. Hydrological models can vary in many ways, including their spatial and temporal resolution, process representation, data requirements, and scale (catchment, national, global, etc). Although there are a number of ways models can be classified, not all models fall into a single category (Singh, 1995). Hydrological models can be grouped generally by their structural classification (empirical, conceptual, physical), or by their spatial discretisation (lumped, semi-distributed, distributed), and are mostly a combination of these classifications. Identifying the main aims and priorities of the modelling study and any limitations in data availability helps with the choice of model, however, it is ultimately a subjective decision by the modeller based on preference and what is task-appropriate (Wagener et al., 2003).

2.1.1. STRUCTURAL CLASSIFICAITON

Hydrological model structures can be classified into three broad groups – empirical, conceptual and physical models. These vary in the way that hydrological processes are represented, and with what degree of complexity, in the model code. These model structures are summarised in Table 2.1 with their advantages, disadvantages, and some examples.

EMPIRICAL MODELS

Empirical models are the simplest of these general structures, which use a simple statistical relationship between inputs and outputs to make rainfall-runoff predictions (Devi et al., 2015). They are sometimes called data-driven (Kokkonen et al., 2001) as they only use information that is available in existing data, without considering characteristics and processes occurring in the hydrological system. For this reason, they are known as 'black box models', as very little is known of the internal processes that are controlling the runoff generation in the catchment (Beven, 2012; Granata et al., 2016). Very few parameters are needed, and this makes empirical models computationally inexpensive and fast to run.

CONCEPTUAL MODELS

Conceptual models aim to describe the dominant components of the hydrological cycle. They consist of several stores that aim to represent the physical landscape of a catchment. Simplified equations of hydrological behaviour and processes govern these models. These are usually versions of the water balance equation, representing the relationship between rainfall to runoff, evaporation and percolation to groundwater (Vaze, 2012). Variables within these equations are interpreted with model parameters, which are determined with a combination of observational data and calibration (Wagener, 2003). The data requirements for calibration is determined by the number of parameters used in the conceptualisation of the catchment. Many conceptual hydrological models have been developed with varying degrees of complexity which is dependent on the equations and parameters chosen to represent the system (Beven, 2012; Pechlivanidis et al., 2011). This variation in sophistication of conceptualisation means these models require a range of hydrological and meteorological input data to calibrate and force outputs. Some examples of conceptual models are TOPMODEL (Beven and Kirkby, 1979), Hydrologiska Byråns Vattenbalansavdelning (HBV; Bergström, 1992) and National Weather Service River Forecast System (NWSRF; Burnash et al., 1973).

Due to the simplicity of conceptual models, they are usually computationally inexpensive and useful in multiple-hypotheses testing and uncertainty estimation studies where large ensembles of model simulations are required (Chun et al., 2009). Model parameter identification is a fundamental challenge that hydrologists face (Sivapalan, 2003; Duan et al., 2006) and an advantage of conceptual models is they can be straightforward to calibrate, but this is dependent on the number of parameters and the data requirements that these demand. They can also provide a benchmark of performance for more sophisticated physical models (Orth et al., 2015) to help determine whether the increased level of complexity is of added value for a particular case (Gutz et al., 2003; Perrin et al., 2006; Kobierska et al., 2013).

However, conceptual models also have disadvantages associated with them. For example, parameters are interconnected within conceptualisations of catchment functioning, and this can make a priori parameter estimation methods difficult to implement and unreliable (Wagener and Wheater, 2006). Also, conceptual models can become overparameterised when calibrated (Perrin et al., 2001; Das et al., 2008; Andressian et al., 2012) and parameters lose their physical realism and therefore cannot be mapped back to catchment characteristics (Wagener, 2003).

PHYSICAL MODELS

In physical models (also called process-based models), hydrological processes are modelled explicitly using known physical laws (Li et al., 2015) and solve equations that express the conservation of mass, momentum and energy (Kampf and Burges, 2007) and represent observable hydrological state variables and fluxes, often from laboratory based theory. Physical models incorporate both physical and process parameters. Physical parameters are used to describe the physical characteristics of the catchment and can be directly measured. Whereas process parameters represent physical properties, for example average water storage capacity (Pechlivanidis et al., 2011) and cannot be directly measured. One of the greatest advantages of physical models is that there is a direct connection with many of the model parameters and physical characteristics of the catchment, and this can make them seem more realistic. Parameters can also be regionalised and can be used to make predictions in ungauged basins that have similar characteristics to gauged catchments (Sivapalan, 2003, Garambois et al., 2015, Hundecha et al., 2016). There are several examples of physical models including Soil and Water Assessment Tool (SWAT; Neitsch et al., 2009), MIKE-SHE (Abbott et al., 1986), Visualising Ecosystem Land Management Assessments (VELMA; Abdelnour et al., 2011) and Penn State Integrated Hydrological Modelling System (PIHM; Qu, 2004).

However, process-based models raise several issues. There has been much debate on how to ‘correctly’ approach process-based modelling (Sivapalan et al., 2003; Maxwell and Miller, 2005; Beven and Cloke, 2012), mainly focusing on how to appropriately parameterise processes, limitations of available (or unavailable altogether) data, and computational expense and constraints of evaluation and analysis (Clark et al., 2017). The physical principals that underpin the model’s formulation are often based on laboratory or small-scale field experiments, and this introduces uncertainty when scaled up. Extrapolation to larger scales assumes that scaling up of processes is linear and unaffected by space. This raises questions about the applicability of this type of model (Beven, 2004). Another issue that is sometimes faced is the use of simplified versions of explicit physical equations in the attempt to increase computational efficiency, for example the use of simplified 1D St. Venant equations (Clark et al., 2015a). Catchment heterogeneities, e.g. soil, geology, land use, etc. also provide an obstacle to the accurate building of a process-based model, as they make model structural identification difficult. Available data is mainly made of point measurements in space (Wheater, 2002) and these are then averaged and extrapolated onto a predefined grid scheme, which inevitably loses spatial variation, no matter how fine the resolution is.

2.1.2. SPATIAL DISCRETISATION

The defining of a hydrological model as being either lumped or distributed refers to the spatial discretisation. Table 2.1 shows a comparison and a summary of the properties associated with broad categories for spatial discretisation. A lumped model treats the catchment as a single unit, and the state variables are averaged over the whole catchment area (Beven, 2001) effectively losing all spatial variability. In contrast, distributed hydrological models produce streamflow predictions that are distributed in space. Catchments are discretised into sub-sections, for example grid squares or hydrological response units, and the model equations are solved for the state variables that are associated with the smaller unit (Singh and Frevert, 2006). Distributed models have the capacity to apply spatially varying data (e.g. high-resolution precipitation, soil, temperature, land use, etc.) as forcing inputs (Carpenter and Georgakakos, 2006). Spatially varying precipitation and catchment characteristics can have significant impacts on the hydrological response of basins (Khakbaz et al., 2009). Characterising and modelling the relationship between spatial distribution of rainfall, basin properties and runoff generating mechanisms has been an ongoing research question for hydrologists. Another advantage of distributed models, as well as the potential to improve discharge simulation at the catchment outlet, is the capability to produce predictions at ungauged locations where measurements are not available (Koren et al., 2004).

Although distributed models can simulate spatial and temporal variations, they often use data that is of a coarser resolution than the grid-square cells within the model, therefore using average values for variables and parameters (Beven, 2001). As a result of this issue, Reed et al. (2004) conducted an inter-comparison study of several lumped and distributed models and concluded that the lumped models outperformed the distributed models in more tests. Given these conclusions and the data requirements to effectively parameterise and validate distributed hydrological models, semi-distributed models have been developed (Abu El-Nasr et al., 2005). They are used to combine the advantages of both lumped and distributed modelling. Instead of attempting to represent the full spatial variability of state variables, semi-distributed models instead discretise the catchment to a degree that is thought to appropriately capture the dominant catchment processes. Therefore, the 'most important' features of a basin can be represented, while requiring less data and lower computational cost than distributed models (Orellana et al., 2008).

Table 2.1. Comparison and summary of basic structures and spatial discretisation for hydrological model

STRUCTURAL CLASSIFICATION				
	Method	**Strengths**	**Weaknesses**	**Examples**
EMPIRICAL	Data driven, look to find relationship between inputs and outputs of rainfall-runoff	Computationally inexpensive, can be run quickly, small number of parameters	Very sensitive to errors and uncertainty in input data, No connection with physical characteristics of catchments	SCS-Curve Number in SWAT, regression equations
CONCEPTUAL	Simplified equations that represent the water balance and storage in catchments	Simplified model structure, can be easy to calibrate, computationally inexpensive	Parameters may not be physical representations of catchment characteristics	HBV, TOPMODEL, NWSRF
PHYSICAL	Physical laws and equations that are based on real and observable hydrological responses	Incorporates spatial and temporal variability, parameters and governing equations are physically realistic	Large number of parameters, large data requirements for calibration, computationally expensive	MIKE-SHE, VIC, SWAT, PIHM, VELMA
SPATIAL DISCRETISATION				
LUMPED	Catchment is treated as a single unit, spatial variability is not included, input data is averaged for catchment	Good if aim is to simulate average conditions in a catchment, fast run times	Not good for large catchments, spatial resolution and variability is lost, assumptions about catchment functioning made	Empirical and conceptual models
SEMI-DISTRIBUTED	Combination of lumped and distributed parameters	More spatial variability and representation, uses both lumped averages and sub-catchment specific data	Some loss of spatial resolution with catchment being divided up into grid cells, HRUs, sub-catchments, etc	Conceptual models, and some physical models HYPE, SWAT
DISTRIBUTED	Spatially variability represented	Results can be mapped back into space, hydrological processes are represented	Require large amounts of data for large number of processes and parameters, computationally intensive	Grid-to-Grid, SHETRAN, VELMA, SWAT, MIKE-SHE, PHIM

2.1.3. HYDROLOGICAL MODEL FORCING DATA

Input data requirements to force hydrological models vary with the choice of model structure and spatial discretisation, but generally, data that may have value to a range of modelling applications depending on the complexity of the approach include:

1. Meteorological inputs to the catchment – e.g. precipitation, solar radiation, temperature.
2. Basin characteristics – e.g. topography, catchment area, land use, geology, land cover, soil type.
3. Losses – e.g. evapotranspiration, infiltration, groundwater.
4. Calibration data – e.g. discharge, groundwater levels, soil moisture content.
5. Human effects – e.g. reservoirs, location of dams, land use changes.

The type of hydrological model chosen for a modelling study is in part dependent on a) the availability and quality of these data and b) the objectives of the modelling exercise and the performance of hydrological models partly depends on the quality of data available for both model set-up and forcing (Beven, 2012). In the past, hydrological modelling studies have been limited by data availability and measurement techniques, however, many of the data collection techniques that are relevant to rainfall-runoff modelling have been improved (Beven, 2012). Hydrological properties and fluxes have typically been measured with in situ monitoring systems (Chen and Han, 2016). Although these are an essential source of data for hydrological research, they do have many disadvantages, such as limited spatial and temporal resolution, inconsistencies in the measurements, and they can be expensive to implement and maintain. The use of satellite-based remote sensing has majorly increased the amount of data available for hydrological modelling studies, such as global precipitation data (Bitew and Gebremichael, 2011; Tian and Peters-Lidard, 2010; Kidd et al., 2011), soil moisture content (Entekhabi et al., 2010; Brown et al., 2013), water storage (Tapley et al., 2004), and climatic conditions, such as temperature and carbon fluxes. These global data sets have been particularly helpful in hydrological model development in data scarce and ungauged basins, where very little, or none at all, hydro-meteorological data exist (Shrestha et al., 2006). However, being able to obtain good model performance using these datasets depends on the resolution of the data and the size of the catchment/model domain (Burlando and Rosso, 2002). The uncertainty associated with hydrological model input data will be discussed in Section 2.3.

2.1.4. LARGE SCALE HYDROLOGICAL MODELLING

The growing concern due to climate change and anthropogenic activities affecting the hydrological cycle has drawn attention to the need to provide information on the present and future hydrological dynamics and physical processes occurring on large spatial scales (Doll, 2009; Hannah et al., 2011). Understanding these processes is fundamental to the development and improvement of large-scale hydrological models and being able to produce accurate simulations and predictions (Yang and Musiake, 2003; Dankers and Feyen, 2008; Doll et al., 2008; Doll and Fiedler, 2008; Hangemann et al., 2008; Pappenberger et al., 2008; Getirana et al., 2010). Large scale hydrology encompasses spatial scales greater than a single catchment all the way to the global-scale (Cloke and Hannah, 2011). Large-scale hydrological research is needed to discern driving forces and patterns in the responses of the water cycle, and to identify regions that are most vulnerable and susceptible to climate change and the intensification of the hydrological system (Cloke and Hannah, 2011). This information will be useful for decision and policy making concerned with water hazards, such as floods and droughts (Huang and Demuth, 2010).

The increasing pressure on water resources has demanded the need for a more national government approach to its management (McMillan et al., 2016). This centralised decision making requires national, and sometimes international depending on the river basin, information on the hydrological cycle. This information is needed at both gauged and ungauged catchments. In order to provide information at ungauged points in a river network, large-scale hydrological models are becoming a useful tool (Booker and Woods, 2014; Archifield et al., 2015). Hydrological modelling at large scales has the potential to include many different catchments and river basins that encompass different climatic and physiographic zones (Alcamo et al., 2003; Raje et al., 2013; Widen-Nilsson et al., 2007). One of the benefits gained from this is improving our understanding of differences in model structures and their suitability for different hydrological regions (Gupta et al., 2014).

Global datasets that quantify the spatial and temporal variability of the water cycle are incredibly important to hydrologists as they have helped develop understanding of the dynamic processes occurring over large domains (Syed et al., 2008; Simmons et al., 2010) and how these vary substantially in space. The growth in availability of global data from satellites and remote sensing in the last two decades (Tang et al., 2009) has facilitated the expansion of large-scale hydrological models (Sood and Smakhtin, 2014), which have previously only been implemented in the developed, data-rich areas of the world. National and global climate and river discharge data are also becoming easier to access due to many open-source data portals available (McMillan et al., 2016), for example the Global Runoff Database (GRDC). However, to improve the predictions made by large-scale hydrological models, global datasets need to continue to be developed and improved, for example increasing the spatial resolution in order to better describe the spatial heterogeneity of the climate variability, e.g. precipitation and evapotranspiration.

There is an increasing number of models that simulate the hydrological cycle on large scales, from national, to whole continents, to the whole globe. These models have been developed for a number of different purposes, for example to quantify the water cycle on different spatial scales, to examine climate change impacts on future water resources, and to assess the impacts and dynamics of different hydrological extremes (Lopez Lopez et al., 2016). Some examples of hydrological models that have been applied at large-scales are VIC (Liang et al., 1994, 1996), WaterGAP (Alcamo et al., 2003), ORCHIDEE (d'Orgeval et al., 2008), HTESSEL (Balsam et al., 2009), PCR-GLOBWB (Van Beck et al., 2011), SUPFEX-TRIP (Decharme et al., 2010, 2013), SWAT (Neitsch et al., 2009), and W3RA (van Dijk, 2010; van Dijk et al., 2014). These models that aim to simulate the hydrological processes on a continental/global scale are similar to small-scale catchment models (Sood and Smakhtin, 2014), but they are different in the way processes are represented in the model equations, how they are parameterised and the input data requirements (Haddeland et al., 2011). There are many examples in the literature of efforts being made to evaluate large-scale hydrological models, for example research focusing on comparing observed and modelled river discharge on the continental scale (e.g. Gerten et al., 2004; Decharne and Douville, 2007; Balsamo et al., 2009; Hagemann et al., 2009). There are also examples of studies that have developed models with high spatial and temporal resolution on national, continental and global scales (e.g. Doll et al., 2003; Hunger and Doll, 2008; Troy et al., 2008; Widen-Milsson et al., 2009; Stahl et al., 2011).

However, inevitably there are a number of challenges encountered when using large-scale models to simulate the hydrological dynamics of a chosen domain. When modelling at a scale larger than a single catchment, model structure and parameters need to be selected to describe a number of catchments with varying processes and characteristics. In response, recent work has focused on the need for the hydrological modelling community to vary model structure based on the dominant properties of a catchment (Clark et al., 2008, 2011; Coxon et al., 2014; Fenicia et al., 2011). However, when building a large-scale hydrological model, the model structure may be simplified with just one used throughout the domain. In this case, a more complex structure may be required to have the capacity to simulate the wide range of catchment characteristics and hydrological responses in the domain (McMillan et al., 2016).

When using a large domain in a modelling study, the physical properties within the catchments (e.g. topography, soil type, land use, geology, etc) have a much larger spatial variation and manifest large heterogeneity in hydrological responses to these characteristics (Sawicz et al., 2011; Coron et al., 2012). This will affect the identifiability of optimal model parameters. On a small scale (up to 100km2 ;Liebscher, 1993) geomorphological properties of a catchment are the main driving factors of hydrological behaviour. However, at the large scale (>10,000km2 ; Liebscher, 1993) hydrological responses are controlled by spatial variability of bio-geophysical characteristics. Also, in large river basins, human activities such as irrigation and hydropower production (Montanari et al., 2013), will have an important influence on river flows.

<u>ANTHROPOGENIC INFLUENCES IN RIVER BASINS</u>

In recent decades, the terrestrial water system has experienced drastic changes due to anthropogenic alterations of land use, land cover, and the management of surface water and groundwater systems (Bondeau et al., 2007; Gerten et al., 2007; Rost et al., 2008). The world's rapidly growing population is increasing food demands, which has seen a drastic expansion of global irrigated areas (Siebert et al., 2015) which are highly water intensive. Extremely rapid urbanisation and economic development in many parts of the world are also key drivers in the increasing demands for water globally (Wada et al., 2016). To meet these demands, humans extract large amounts of water from surface and groundwater supplies (Siebert et al., 2016; Siebert and Doll, 2010; Wisser et al., 2010; Konikow, 2011), and these abstraction volumes have increased almost 8-fold in the last 100 years (Oki and Kanae, 2006; Hoekstra and Chapagain 2007; Hanasaki et al., 2008a, b; Wada et al., 2014). Thousands of dams have been built in river basins across the world (Lehner et al., 2011), and these have been used to increase water supply, assist with flood monitoring schemes, and as a source of hydroelectric power generation (Liu et al., 2015, 2016). This unprecedented growth in human demands on water supplies has significantly modified many hydrological processes and at various spatial scales (Sivapalan et al., 2012; Sivapalan, 2015).

Most large-scale hydrological modelling studies on changes in the hydrological cycle focus on impacts of a changing climate and climate extremes (Hirabayashi et al., 2013; Orlowsky and Seneviratne, 2013; Dankers et al., 2014; Jongman et al., 2014; Prudhomme et al., 2014), however, human water management is also an important influence affecting water supply and hydrological variability, from the catchment scale, to national and global scale (van Loon et al., 2016; Di Baldassarre et al., 2017). Recently, modelling studies have focused on human interventions in the hydrological cycle explicitly and this facilitates the attribution of hydrological extremes to either natural/climatic or anthropogenic processes (van Dijk et al., 2013; van Loon and van Lanen, 2013; Veldkamp et al., 2015; He et al., 2017). Hydrological modelling studies that have incorporated anthropogenic influences on catchment processes usually use fairly simplistic representations of these impacts (Wada et al., 2017). This is due to the high uncertainty in human behaviour, a lack of data to parameterise these factors (Veldkamp et al., 2018) and the complex interaction between human activities, climatic characteristics and the hydrological processes in the region (Uhlenbrook et al., 2003).

2.2. EVALUATION OF HYDROLOGICAL MODELS

A 'perfect' hydrological model cannot exist, due to the inability to represent all the physical processes occurring in a catchment, with the complexity of spatio-temporal variability, and therefore all models are in error (Freer et al., 2004). Hence, a large issue in hydrological modelling is being able to discriminate between model representations and deciding which is most appropriate for the task (Dunn et al., 2008). It is important to be able to assess the performance of different models as hypotheses of the real-world system, especially as many different structures may be considered adequate representations of the catchments of interest (Beven and Freer, 2001b). A crucial aspect of hydrological modelling is implementing robust model performance tests.

2.2.1. OBJECTIVE FUNCTIONS

An objective function, or sometimes described as a 'goodness of fit', is a numerical measure of the difference between the simulated model outputs and the observed data for a catchment (e.g. Schaefli and Gupta, 2007). There are many objective functions that can be used by a modeller to assess the performance of simulations, but this is a difficult task for a number of reasons (Pushapalatha et al., 2012). 1) river discharge vary by orders of magnitude, and not all of this information may be of use to the modeller as their study may be focusing on a certain type of flow, such as high flows or low flows. 2) Errors produced by hydrological models are often heteroscedastic, meaning that their variance is not equal across the range of values. 3) The range of discharge values that are being targeted in the performance tests may vary significantly between the different periods of evaluation. 4) The choice of performance criteria may be dependent on the different applications of the model.

For these reasons, a large variety of objective functions have been developed in the literature, as is seen by the review of hydrological modelling performance measures by Dawson et al. (2007), Moriasi et al. (2007) and Reusser et al. (2009). Within this variety of criteria, there are two main categories: absolute criteria, such as Root Mean Square Error, and relative criteria, which is normalised, such as Nash Sutcliffe Efficiency (NSE, Nash and Sutcliffe, 1970). NSE has received much attention in the literature and is a popular performance metric in hydrological modelling studies (e.g. McCuen et al., 2006; Schaefli and Gupta, 2007; Clarke, 2008; Gupta et al., 2009; Moussa, 2010; Gupta and Kling, 2011). NSE compares model errors with the errors in a reference/benchmark model often against the mean flow (Seibert, 2011; Perrin et al., 2006). This provides useful information regarding model simulation performance as it indicates whether the model is performing better or worse than the benchmark. NSE is also dimensionless, and so can be used to compare results between catchments. However, there are a number of weaknesses of the NSE metric, for example, the emphasis on high flows, sensitivity to the hydrological regime, sample size and outliers (Pushapalatha et al., 2012).

2.2.2. MULTI-OBJECTIVE EVALUATION

There has been a shift in hydrological research away from the use of a single objective function, to using multi-objective approaches to evaluate model performance (e.g. Wagener et al., 2005; Winsemius et al., 2009; Ritter et al., 2013; Hrachowitz et al., 2014; Beck et al., 2016). In practical applications of hydrological models, for example real-time flood forecasting, there may be a requirement to accurately simulate the entire hydrograph (Moussa and Chahinian, 2009), therefore, using a single objective function may not exploit all the information in the output data (Wagener, 2003). A hydrological model’s outputs are strongly influenced by the choice of objective function (Abbaspour et al., 2007; Park et al., 2014), and so using a single objective function introduces error and bias into the model predictions. Multi-objective approaches can address this problem.

2.2.3. HYDROLOGICAL SIGNATURES

Within hydrological modelling research, there has been a move away from the traditional use of statistical measures of model performance towards time-step based performance measures and hydrological signatures (Gupta et al., 2008; Yadav et al., 2007; Yilmaz et al., 2008; Krueger et al., 2010). Hydrological signatures are values derived from observed or modelled hydrological data, for example precipitation, discharge, or soil moisture (McMillan et al., 2017) and aim to provide an insight into the functioning of a catchment (Sawicz et al., 2011). They are designed to maximise the information that can be gained from the data, to identify dominant processes occurring in a catchment, and the spatial and temporal variability in rainfall-runoff generating mechanisms (McMillan et al., 2017). Examples of typical signatures that are calculated in the literature from river discharge include mean flow, baseflow index and the slope of the flow duration curve (FDC).

The confidence that we have in predictions made by hydrological models will depend on their capacity to simulate observed data (Westerberg et al., 2011). Traditional statistical methods of model evaluation can be problematic for model structural identification for a number of reasons: 1) the uncertainty in the observation with which model simulations is being compared, 2) the sensitivity of performance measures to different flow magnitudes, 3) the influence of natural uncertainty, that arises from our imperfect knowledge and understanding of environmental processes, that cannot be accounted for, and 4) where the observational data does not overlap with the time period of the model predictions, model performance cannot be evaluated. Hydrological signatures have been proposed to address these problems. Signatures have been used in hydrological analyses for a long while, but it was the concept of hydrological signatures that was first explicitly introduced by Gupta et al. (2008) in the context of targeting relevant information in hydrological datasets for the purpose of model evaluation.

Most studies will use a combination of different signature indices in order to reflect different types of catchment response and behaviour, for example being able to evaluate high and low flows well, and also the flow timing and time to peak. Therefore, using hydrological signatures can help to advance our understanding of the relationship between models and the physical hydrological cycle (e.g. Gupta et al., 2008; Yilmaz et al., 2008; Hingray et al., 2010; Wagener and Montanari, 2011). The use of hydrological signatures has both advantages and disadvantages. One of the main advantages is that signature indices are physically interpretable, as they represent the physical processes (Euser et al., 2013). Therefore, this should make the results more transferable to catchments with similar properties and characteristics. However, the disadvantage of signature indices is that they are designed to represent a specific type of catchment behaviour or response, and this is at the expense of others, and therefore it is necessary to consider a variety of signatures in order to gain more information about the catchment's functioning (Euser et al., 2013).

2.3. HYDROLOGICAL MODEL UNCERTAINTY

Uncertainty analysis is the quantification and calculation of the confidence in a model's simulations being an accurate representation of real world physical processes (e.g. Singh, 1995), and aims to examine the combined effects of specific sources of uncertainties. Uncertainty analysis is essential in any modelling study where simulations are being used for informing decision making, hypothesis testing of catchment functioning, and highlighting aspects of models that require further development or improvement. However, calculating uncertainty can prove to be a difficult aspect of modelling, as it combines the errors in: 1) observations and input data due to variability of spatio-temporal scales, 2) model parameters, 3) the equations chosen to represent the hydrological processes, and 4) boundary conditions (Salamon and Feyen, 2010).

Uncertainty can be divided into two main categories. These are aleatory and epistemic uncertainties (Merz and Thieken, 2005). Aleatory uncertainty arises due to the natural variability in a system and cannot be reduced (Gong et al., 2013). Epistemic uncertainty is due to our incomplete knowledge of how environmental systems work, and the different processes and intricacies that are associated (Apel et al., 2008). This uncertainty can be reduced by improving the quality of measurement data, using more detailed experimental field studies to enhance our understanding of hydrological processes, and using hydrological models to test multiple hypotheses about the functioning of the water cycle. Modellers are faced with a difficult task when choosing what they believe to be an appropriate model for their study, due to these quantifiable and unquantifiable uncertainties (Lee et al., 2011).

2.3.1 SOURCES OF UNCERTAINTY IN HYDROLOGICAL MODELLING

In any robust and meaningful hydrological research, an assessment and understanding of the sources of uncertainty is fundamental (Pechilivanidis et al., 2011). Uncertainty in modelling hydrological systems has been summarised with three primary sources by Di Baldassarre and Montanari (2009): a. uncertainty in observations, b. parametric uncertainty, and c. model structural uncertainty.

2.3.2. DATA UNCERTAINTY

Uncertainty is present in all data as all observations and measurements are subject to error. Errors can be classified as systematic, random, or spurious. Some examples of errors that can arise in the data, and the uncertainty that they cause, that are used in hydrological models are: 1) the discrete time and point measurement nature of data and the way in which it is collected gives little insight into the variation that may be occurring between measurements, 2) Interpolation of point measurements in space loses the spatial variability, 3) observations and measurements may not be directly related to the process that is being estimated, e.g. using air temperature in the estimation of potential evapotranspiration. These errors in turn affect the estimation of model parameters and choice of structure (Kavetski and Clark, 2010). Rainfall, PET and discharge data uncertainties are specific to hydrological modelling studies.

RAINFALL DATA UNCERTAINTY

Precipitation is one of the key drivers of the hydrological cycle, and is a critical aspect in any hydrological modelling study (Kidd and Huffman, 2011; Hou et al., 2014). Accurate representation of rainfall in models is needed to produce accurate simulations of streamflow (Beven, 2004), and this plays an important role in the management of water resources and predicting hydrological extremes, such as flood events and drought. Generally, precipitation data is obtained using two methods: ground-based measurements, and datasets from satellite and remote sensing observations (Alijanian et al., 2017). Ground-based observations are in the form of rain gauges and weather radars. These are relatively straightforward to implement, for example using a tipping bucket methodology. Ground measurements of precipitation are usually used to represent regions with a size of 10-100km2 (Chao et al., 2018). However, rain gauge measurements are site-specific, and gauge networks are often sparse and uneven in space, which makes this source of precipitation data unrepresentative (Strauch et al., 2017). These measurements often have to be extrapolated, and this does not represent the spatial variability of the rainfall across a catchment (Villarini et al., 2008).

There has been a rapid development in satellite and remote sensing based observations of precipitation, and the creation of global open-source datasets has facilitated the advancement of hydrological modelling in the data scarce regions of the world. Satellite-based datasets have the ability to overcome some of the issues that are presented by point-specific gauge measurements of rainfall, for example, they have the capacity to cover the whole globe and produce a large amount of data (Kidd et al., 2011). These global satellite-based datasets are often in the form of a global grid, with fairly coarse spatial resolution, and therefore there are inevitably large errors and uncertainty in these datasets. Precipitation patterns have an ‘intrinsic irregularity’ (Molini et al., 2001) and this makes it very difficult to measure the real time and space evolution of precipitation.

POTENTIAL EVAPOTRANSPIRATION UNCERTAINTY

Potential evapotranspiration (PET) is defined as the rate at which evapotranspiration (ET) would occur if an area was uniformly covered with vegetation and had access to a consistently sufficient water source, and actual evapotranspiration (AET) is the actual rate of evapotranspiration occurring from the land surface (McVicar et al., 2012, McMahon et al., 2013; Li et al., 2016). PET is an input in hydrological models, and defines the upper limit of AET. Evapotranspiration (ET) transfers a large amount of water from land surface into the atmosphere, which is also closely associated with land-surface atmosphere exchanges of carbon and energy (Van Camp et al., 2016). It is one of the most important fluxes in the hydrological cycle, and therefore accurate representations are crucial for making predictions of discharge in catchments (Badgley et al., 2015). ET includes the evaporative fluxes from soil and vegetation, and transpiration fluxes from vegetation and the canopy (Montheith, 1965; Shukla and Mintz, 1981). 85% of the global ET total is from transpiration fluxes from the canopy, and this returns around 50% of precipitation back to the atmosphere (Oki and Kanae, 2006). In the arid and semiarid climatic zones, ET can be responsible for returning more than 95% of the annual precipitation (Kurc and Small, 2004).

ET is one of the most difficult meteorological components of the water balance to measure and estimate (Lettenmaier and Famigletti, 2006). Direct measurements of ET can be made using methods such as flux towers (Shi et al., 2008; Baldocchi and Ryu, 2011). However, these are only available at small scales and mostly only available in developed countries (Miralles et al., 2016). Numerous approaches have been developed to retrieve ET on large scales from satellite observations (Dugo and Gao, 2011; Kalma et al., 2008; Liou and Kar, 2014; Mercado et al., 2009; Miralles et al., 2016; Su et al., 2010). There have been many models established that estimate evapotranspiration from this satellite data, for example Pennman-Monteith (Monteith, 1965), Priestly-Taylor (Priestly and Taylor, 1972), Hargreaves (Hargreaves and Samani, 1985). These equations require information that can be grouped into three broad categories – net radiations, meteorology (e.g. water vapour pressure, air temperature, and wind speed), and vegetation (e.g. land cover, vegetation greenness indices, and leaf area index (LAI) (Mu et al., 2007, 2011; Fisher et al. 2008). Uncertainty in PET datasets arises due to the large range of models used to produce these estimates for large scales. A number of remotely sensed data is required as inputs into these algorithms, and therefore the uncertainty of PET estimates is a combination of the errors and uncertainty in these inputs.

DISCHARGE DATA UNCERTAINTY

In hydrology, stream flow time series are commonly derived from stream level heights and a rating curve that describes the relationship between stage and discharge. This is a key uncertainty in hydrological studies as discharge observations are often the only available data to evaluate models against (e.g. McMillan et al., 2010; Jalbert et al., 2011; Domeneghetti et al., 2012; McMillan and Westerberg, 2015). Rating curve uncertainty occurs due to the assumption made that there is a discrete relationship between stream stage and discharge gaugings (Tomkins, 2012). However, during extreme out-of-bank flows channel shapes can be drastically altered, vegetation growth, and erosion/sedimentation at the gauging sites can all change streamflow (Di Baldassarre and Montanari, 2009). The fitting of the rating curve to the stage and discharge measurements introduces additional uncertainty, especially where the curve is based on a limited number of observations, or where there is a large scatter in gauge measurements (Tomkins, 2012). Further uncertainty is identified here also, as gauging measurements themselves are uncertain due to imperfect measurements, are subject to variability which can impact and change the stage-discharge relationship. Rating curve uncertainty has been the focus of much hydrological research (e.g. Fenton and Keller, 2001; Moyeed and Clarke, 2005; Petersen-Overleir, 2006; Dottori et al., 2009; Reitan and Petersen-Overleir, 2009; Westerberg et al., 2011; Coxon et al., 2015).

Model parameter and structure selection is affected by this uncertainty in the stage-discharge relationship, as discharge measurements are used to calibrate and evaluate model simulations (McMillan et al., 2010). Parameter identification is also affected by other input data errors, for example, uncertain rainfall data can have a significant impact on model performance when simulations are compared with observations, and also make it difficult to select an 'optimal' parameter set (Pappenberger et al., 2005; Oudin et al., 2006; Arnaud et al., 2011). However, input data uncertainty does not necessarily always have a negative impact on model performance, as it can be compensated for during model parameter calibration (Gourley and Vieux, 2006; Liu et al., 2009).

2.3.4. PARAMETRIC UNCERTAINTY

Inevitably, all data is subject to some error and uncertainty, and this is translated into uncertainty and bias in parameter estimation (McIntyre et al., 2002; Freer et al., 2004). The majority of hydrological models are conceptual, and therefore obtain parameter values through calibration, by searching a parameter space to find an 'optimal' set. As discussed in earlier sections, this can be problematic, as it is sometimes difficult to find unique and physically realistic parameters. Calibration can also be heavily biased, due to the influence of the chosen objective function with which to assess the model performance (Hunter et al., 2005, 2007).

When calibrating a hydrological model, caution is needed to avoid over-parameterisation, and leading to equifinality (Beven and Freer, 2001b), meaning that more than one parameter set/model structure is equally capable of producing an acceptable representation of the observed processes. Beven (2006) questions the existence of an 'optimal' parameter set. This is supported by Bates et al. (2013) who say that there is a large degree of freedom in hydrological modelling and therefore it is likely that a range of different parameters will be able to fit observations equally well. The Monte Carlo method is an example of a framework that has been developed to address the uncertainty regarding parameter estimation. This method randomly samples parameters from a probability distribution in the search for an optimal set when evaluated using a chosen performance measure.

2.3.5. STRUCTURAL UNCERTAINTY

Structural uncertainty is an inherent source of error in hydrological modelling (Beven and Binley, 1992). Structural uncertainty can be described in terms of its insufficient performance and non-uniqueness (Hublart et al., 2015). Structural insufficiency is due to the assumptions that are made when the model’s governing equations are chosen, and which processes, and how to represent them, are defined. Non-uniqueness describes the existence of many model structures (and parameterisations) and can produce equally acceptable simulations when evaluated against the observational data (Beven, 2006). This has led to some researchers in the hydrological community to question the traditional approach to hydrological modelling of using a 'one-size-fits-all' structure (Savenije, 2009) and whether this is able to capture the diversity of a system. This method can prove to be problematic when choosing a model structure to use in the study. The modeller has many options during the model build stage, and must decide what processes, and what level of complexity to represent these with, in the model code (Renard et al., 2010). However, this is usually based on the modeller's preferences and what they deem to be most purpose-appropriate, and this introduces bias and error into the results (Krueger et al., 2010).

A fixed model structure that is applied to all catchments in a study does have many attractive features, including 1) it may be computationally more efficient than applying many models of varying complexity and input data demands, 2) the repeated application of a single model facilitates it's further development and improvement, and advances our understanding of catchment characteristics and model dynamics, 3) a single model structure also makes it simpler to identify the relationship between catchment characteristics and parameters, benefiting regionalisation studies, and 4) makes comparison of simulations between catchments simpler, as it reduces the uncertainty relating to differences that arise due to behaviour of different models. However, there are many practical weaknesses of using a single model structure. One of these limitations that is seen in many hydrological modelling studies is the need to incorporate specialised model components for specific catchment properties, that can be turned on and off in the model code. For example, in many models the representation of snowmelt is usually required as an additional module, which implies that certain catchments need different model structures in order to capture the climatic influence on the hydrological response of a catchment. For urban catchments a representation of impervious regions is sometimes needed (Cuo et al., 2008), and in catchments that are highly influenced by geology and groundwater exchange, a component that represents these subsurface fluxes is needed (Le Moine et al., 2007).

Much recent research has focused on investigating uncertainty relating to model structure (e.g. Butts et al., 2004; Renard et al., 2011; van Esse et al., 2013; Moges et al., 2016; Tyralla et al., 2016). Studies have shown that model structural uncertainty can have a strong influence on model performance and can be of a similar magnitude of error as that caused by parameterisation and input data. Wagener and Gupta (2005) found that structural error can be the most significant source of uncertainty within hydrological modelling, emphasising the need to account for this in uncertainty analysis. This research has led to a shift in the hydrological modelling community from selecting a single model structure to using multi-model ensembles (Velaquez et al., 2010; Gudmundsson et al., 2012), multiple model structures of increasing complexity (Farmer et al., 2003; Bai et al., 2009; Pushpalatha et al., 2011) and flexible modelling frameworks (Clark et al., 2008, Coxon et al., 2014).

Model structural uncertainty is usually assessed and identified by investigating the catchment characteristics and behaviour, such a peak discharge, time to peak and runoff volume (Butts et al., 2004).

CATCHMENT CHARACTERISTICS AND PROCESS REPRESENTATION IN MODEL STRUCTURE

'Catchment classification' is a theoretical framework with which to characterise catchments by their variability in 'space, time and process' (McDonnell and Woods, 2004). Understanding the key properties that control similar hydrological behaviour and being able to attribute this to different catchment characteristics is an important task that hydrologists are faced with. Research efforts for catchment classification have mainly focused on parameter regionalisation rather than variability that could be caused by differences in model structure. However, it is highly likely that error associated with structural choices will affect the relationship between calibrated parameter values and catchment characteristics. For this reason, catchments with varying hydrological behaviour associated with different characteristics may require different, or even multiple, model structures to represent this (Gupta et al., 2014).

Dominant runoff-generating mechanisms vary between catchments due to their different properties, for example geology, topography, land use, climate, etc. Therefore, it is intuitive that using different model structures in a catchment will produce varying model performance (Van Dijk, 2010). There are studies that have investigated the degree of structural uncertainty of hydrological models in different catchments. Some studies have found that there seems to be no relationship between dominant catchment characteristics and the choice of model structure (e.g. Perrin et al., 2001; Lee et al., 2005; Hollander et al., 2009), and supporting the 'equifinality' theory (Beven, 2006). However, other studies have found contrasting results, and concluded that catchments with varying characteristics require different model structures to represent the dominant processes. It has also been shown that variations in model performance can be correlated with the different hydrological behaviour of catchments. For example, Fenicia et al. (2014) attributed the difference in model performance for headwater catchments in Luxembourg to geology as being the dominant characteristic control on hydrological processes and responses. Also, Buytaert and Beven (2011) used different models with varying structures to attempt to model catchment processes in upland catchments in Ecuadorian Andes and found that model performance varied depending on the processes that were represented in the structure and the particular catchment of interest.

MULTI-MODEL ENSEMBLES

There has been a great research effort in the development of more hydrological models, but the single 'perfect' model has yet to be built that can be used for all types of application and under all conditions (Smith et al., 2004; Beven, 2006; Duan et al., 2006). This is likely an impossible task, and a futile objective in the design and build of a hydrological model. Different types of hydrological models have strengths and weaknesses in terms of their capacity to represent aspects of the physical processes occurring within a basin. Using a single model can often lead to simulations and predictions that capture some types of hydrological behaviour at the expense of others. A common method used in hydrological research to test understanding of model structural differences is to use model intercomparison experiments. Ensemble approaches to hydrological modelling using multiple parameter sets and structures can help to improve the uncertainty analysis (Vrugt et al., 2003; McEnery et al., 2005). Multiple model structures of varying complexity are forced with the same input data with identical boundary conditions (Reed et al., 2004), and simulations are compared against each other using a choice of performance measure.

Multiple model intercomparison experiments have directed attention of the research community to the range of simulations that arise when different model structures are forced with the same input data, however, they have been less successful in advancing our knowledge and understanding of the reasons for these intermodal differences (Clark et al., 2008). Instead, what is gained from these studies is the different types of hydrological behaviour that can be simulated from the differences in model structures used (Slater et al., 2001).

FLEXIBLE MODEL STRUCTURES

Hydrological model development has largely been in favour of a 'one-size-fits-all' approach, which aims to find one model structure that is applicable to all catchments (Fenicia et al., 2011). However, it seems unlikely that the fundamental hydrological processes are the same in all catchments globally. Therefore, there has been a growing interest in flexible model structure frameworks, which give the modeller the option of adapting the model, and processes that are represented, to the catchment of interest.

Flexible structure methodologies use many different model structures and components (Clark et al., 2011) to assess parametric, structural and data uncertainties (Wagener et al., 2001; Krueger et al., 2010; Smith and Marshall, 2010) analyse how structural decisions affect the representation of hydrological behaviour (Staudinger et al., 2011) and determine what controls the choice of structure regarding catchment characteristics (Lee et al., 2005). Due to considerable uncertainty, and a lack of unified theories of hydrology at a catchment scale (which has been observed by commentators, e.g. Sivapalan, 2005; McDonnell et al., 2007; Troch et al., 2009; Clark et al., 2011), there are many combinations of model structure and parameter sets that can be considered hypotheses of the real world (Buytraert and Beven, 2011). Flexible model structures have the capacity to compare multiple parameterisations and correlate the differences in model performance with specific components of the model structure.

There are a number of examples of existing flexible hydrological modelling frameworks in the literature. One example is the Framework for Understanding Structural Errors (FUSE; Clark et al., 2008). Using four 'parent' lumped conceptual models, the modeller can build multiple hydrological models using different combinations of processes from these, and therefore can test different hypotheses about catchment functioning and behaviour. In the application of FUSE by Clark et al. (2008), the model's structure was found to be of the same importance as the parametrisation. Another example is the Rainfall-Runoff Modelling Toolbox (Wagener et al., 2002). This framework was developed for the identification of parsimonious, lumped model structures. It’s based on a modular structure, consisting of a moisture accounting module and a routing module, and different approaches can be taken to represent these modules. These different representations can be used to test different hypotheses of catchment functioning. SUPERFLEX (Fenicia et al., 2011) is a flexible framework that is based on generic ‘building blocks’ that represent conceptual stores in the model domain. These components are a reservoir element, a lag function element, and a junction element. These are generalised, and within the modelling framework can be arranged to represent different flow configurations and alternative hypotheses of catchment functioning. The Structure for Unifying Modelling Alternatives (SUMMA, Clark et al., 2015b) is based on a set of conservation of mass and energy equations which are the structural core of the modelling framework. SUMMA’s model domain covers the atmospheric zone above the vegetation canopy through to the river channel, and includes the dominant biophysical and hydrologic processes of catchments. Within this framework, the different physical processes occurring in catchments can be represented in different ways, and these can be organised in different spatial configurations that represent different hydrologic connectivities in the landscape.

Flexible structure modelling studies highlight the crucial research question within hydrology of advancing our knowledge and understanding of the complexity of the relationship between model structure, catchment properties and associated hydrological behaviour.

2.4. LARGE SCALE HYDROLOGICAL MODELLING IN DATA SCARCE REGIONS

There are a number of challenges that hydrologists face when attempting to model at large-scales in data-scarce regions of the world. The heterogeneity of catchment properties across large model domains is difficult to account for in model structures and parameter sets. At large scales, the hydrological processes occurring are of increasing complexity, and building a hydrological model that can capture this, and maintain computational efficiency, is difficult. Data-scarcity means that these heterogeneities in factors such as soils, land use, geology, etc. cannot be characterised in model structures to represent many of the world’s river basins (Arnell, 1999; Doll and Siebert, 2002; Fekete et al., 2004; Decharme and Douville, 2006; Guntner, 2008; Hunger and Doll, 2008; Peel et al., 2010; Widen-Nilsson et al., 2009). Therefore, hydrological modelling studies are needed in these data scarce regions of the world in order to test hypotheses of catchment functioning using the limited available data and knowledge to inform decisions.

The Niger basin in West Africa is a large river basin (third largest on the African continent) with very few ground observational data. It has a diverse hydro-climatic regime across the basin, with a very distinct seasonal cycle. There are also large inter-annual variabilities in precipitation and river discharges. This makes the Niger basin an ideal location to evaluate model performance in large and data scarce locations, and to test hypotheses of hydrological behaviour, and the relationship this has with model structural choices and parameterisations.

2.4.1. HYDROLOGICAL MODELLING IN THE NIGER BASIN

THE NIGER RIVER

With a surface area of 2.27 million km2 shared by ten countries, and a population of around 100 million people (George Golitzen et al., 2005), the Niger River basin is one of the most important river basins in Africa, with many of the Western African economies depending on it. The river supports many important uses, such as irrigation, fisheries, proving drinking water, and generating hydropower (Zwarts et al., 2005; Zwarts, 2010). As the Niger Basin spans such a large area, it also encompasses diverse environmental and climatic areas. Precipitation varies across the basin from approximately 2100 mm/yr in the headwaters in Guinea, to around 250mm/yr in the northern regions in the Sahelian belt (Thompson et al., 2016). Potential evapotranspiration is also highly variable, ranging from around 1800mm/yr in the headwaters, to around 2150mm/year in the north (Thomspon et al., 2017). River flows in the Niger have a distinct seasonal regime, with very low flows during the dry season compared to the very high flows that characterise the wet season (Zwarts et al., 2005). The countries that are located in the Niger River basin are also particularly vulnerable to climate change (Amadou et al., 2014; Aich et al., 2016) and the intensification of the hydrological cycle. This is a great concern for the countries that are located in the Sahel, as they are extremely dependant on the water from the river for their economies (George Golitzen, 2005). All of these features make this river of interest to hydrologist to study, both in terms of its dynamic hydro-climatic variability and for the importance of water resources management.

MAIN FINDINGS AND RESULTS IN PREVIOUS WORK RELEVANT TO THIS STUDY

There are a number of examples of studies that have successfully applied different hydrological models to the Niger basin. A range of hydrological models have been used, and the main objective of these studies varies across the literature from predictions of discharge to climate change impact on future flooding. Some examples of these modelling studies are summarised in Table 2.2.

One of the main similarities between the studies that are summarised is that the hydrological models used do well to simulate the dynamics of the observed discharge, but find difficulty when predicting the magnitude of peak flows. The conclusion that is drawn is that there are missing processes that are important to the Niger basin in the model structures. One of the processes that is assumed to be underpredicted in many of these modelling studies is evapotranspiration (Schuol and Abbaspour, 2006; Dadson et al., 2010; Andersson et al., 2017a). In the studies where model performance has been improved, the model structure and representation of processes has been modified to include hydrological characteristics of the Upper Niger. For example, flood inundation modules have been added to represent the Inner Niger Delta, a large seasonal wetland at the downstream region of the Upper Niger. In high flood years, this floodplain can be as large as 40,000km2, and this large area of surface water is an source of a large amount of additional evaporation in this region.

Another common feature of these hydrological modelling studies is that in all but a few, model simulations are produced with a monthly time step. Monthly time steps are usually easier to reproduce 'good' model simulations when evaluated against the observed data (e.g. Dadson et al. (2010) achieved NSE scores of 0.7; Huang et al. (2017) obtained satisfactory results, with thresholds in NSE of 0.7; Thompson et al. (2017b) achieving 'very good' or 'excellent' model performance defined by NSE). This is due to the smoothing of the daily dynamics occurring in the catchment. However, being able to predict daily discharge is an important aspect of hydrological modelling, as these predictions are used in many practical applications, such as flood and drought risk assessments and operational forecasting.

Most of the modelling studies in the Niger summarised in Table 2.2 look just at peak river flows. This is common in hydrological modelling studies, however, being able to produce simulations of low river flows is just as important. Low flow periods and droughts have a drastic effect on the availability of water supplies for many uses, such as irrigation, drinking water, hydro-electric power generation, and ecosystem maintenance (Staudinger et al., 2011). Also, being able to produce both high and low flows with a singular hydrological model is useful for practical applications, as only one model needs to be set up and evaluated. It is also a more powerful test of model performance when assessing capacity to model low and high flows.

This also links with another common feature of these studies in the Niger basin – model performance is only evaluated with one performance metric. All studies use Nash-Sutcliffe Efficiency (NSE) as the objective function with which to evaluate model simulations against observed discharge. This is a widely used metric in the literature to assess the predictive power of hydrological models, as it is simple to implement and allows for comparability across studies. However, the use of this as a singular performance evaluation measure has been criticised (e.g. Oudin et al., 2006; Schaefli and Gupta, 2007). NSE scores are sensitive to the extremes in data and outliers (McCuen et al., 2006), and therefore is able to identify simulations that are good at capturing high flows. In order to produce a more robust evaluation of a model's skill at predicting river flow, a multi-metric criteria may be required (Wagener and Gupta, 2005; Winsemius et al., 2009; Ritter et al., 2013; Hrachowitz et al., 2014; Beck et al., 2016).

In this study, a flexible hydrological modelling framework is applied to the Upper Niger basin. Simulations are produced and evaluated at a daily time-step and a multi-metric evaluation criteria is used to assess the model's ability to reproduce different aspects of flow. Multiple global datasets are used to investigate the uncertainty of these as inputs for hydrological modelling studies, and the effect that their associated errors have on simulated river flow. The model structure was adjusted to test hypotheses of basin functioning, and evaluate the performance when different dominant hydrological processes are represented differently.

Table 2.2. Summary of selected hydrological modelling studies in the Upper Niger basin and the main findings and results.

Reference	Study objectives	Model	Input data	Timestep and period	Main findings/results
Dadson et al., 2010	1. Add inundation module within land-surface model 2. Evaluate model over Inner Niger Delta 3. quantify effects of wetland inundation as a land-atmosphere feedback mechanism	JULES with overbank flow paramerterisation, 0.5 degree regular lat-long grid	ALMIP experiment data	Monthly, 2002-2006	Model managed to capture main features of observed flow and inundation patterns. However, overprediction in modelled discharge by 41%, suggested due to human abstractions and underestimation of evaporation. Model efficiency defined by NSE score was 0.70.
Pedinotti et al., 2012	Evaluate the ability of the ISBA-TRIP continental hydrological model to represent the key hydrological characteristics of the Niger Basin	ISBA-TRIP Continental Hydrologic System, uses flooding scheme and linear deep aquifer reservoir, 0.5 degree spatial resolution	TRMM-3B42 rainfall and RFE2	Daily, 2002-2007	Model provides a good estimate of the surface water dynamics, even with a relatively simplistic channel routing scheme. Flooding scheme increases evaporative model losses, reducing discharge downstream from Inner Delta. However, model struggle to simulate low flows and the annual peaks.
Huang et al., 2017	Multi-model ensemble of 9 regional hydrological models in 12-large-scale river basins, including the Niger. Individual models and ensemble mean run	ECOMAG, HBV, HYMOD, HYPE, mHM, SWAT, SWIM, VIC, WaterGAP3	WATCH forcing data	Monthly, 1971-2001	None of the hydrological models were able to provide satisfactory results for the Niger basin (when models were evaluated using NSE, with a threshold of 0.7). The simulated extreme low flows had a larger bias compared to high flows. Overestimation in flood peaks in the Niger basin for most of the applied model.
Andersson et al., 2017a, b	a. test a framework for improving process-oriented hydrological models that are applied to another region that was first developed for, i.e. HYPE for the Niger basin, when originally developed for Sweden. b. used improved model to estimate peak river flow statistics and test model capacity for forecasting predictions.	HYPE, divides basin into smaller sub-basins	WATCH forcing data	Daily, 1979-2009	Niger-HYPE1.0: evaluated and was found to have inadequately described physical processes. Model originally developed for Sweden. Could not simulate magnitude of daily flow dynamics. NSE -1 in all sub-basins. Niger-HYPE2.0: process refinements for evaporation, flood and river-atmosphere exchanges. Model simulated peak discharge reasonably well, average overprediction of 20%.

Thompson et al., 2017	Assessing future river flows and flood extents in the Upper Niger and Inner Niger Delta under a changing climate. Evaluates hydrological model for input to GCM for uncertainty analysis due to climate change using CMIP5 ensemble.	A semi-distributed conceptual hydrological model developed for Upper Niger, included inundation module. Sub-models with identical structure for 11 sub-basins defined by downstream gauge.	CRU TS 3.0 gridded data	Monthly, 1950-2000	Model includes inundation extents, and module that represents Selingue, Dam and Markala Barrage. Model found to perform well for the calibration period (1950-1976), NSE for validation period (1961-1990) being 'excellent' or 'very good'. Overestimation of seasonal peak discharged common across the simulations. Suggest the reason for this may be the changes in land use in Sahelian region in the recent decades
Schuol and Abbaspour, 2006	West Africa selected as case study for analysis of a large-scale hydrological SWAT model. Aim to address some calibration and uncertainty issues when using SWAT at large scales. Part of case study in larger project to quantify global availability of freshwater.	SWAT, catchment dived into sub-basins	All data taken from global databases	Monthly, 1971-1995	NSE results for the upper Niger for calibration and validation periods were between 0 and 0.7. Conclude that not all processes that are important for the Niger in model. Processes concluded as being large reservoirs and Inner Niger Delta delaying runoff and contributing to higher evaporative losses.
Dezetter et al., 2008	Investigation to find the best combination of hydrological model and data for catchments in West Africa including Niger basin. Used 3 PET grids, and 4 soil water holding capacity grid.	Semi-distributed model developed comprising 2 models: GR2M and WBM, 0.5 degree resolution	CRU gridded precipitation, three PET grid generated from FAO and CRu	Monthly, 1902-1995	Analysis shows models not very sensitive to different PET grids, but much more sensitive to soil grids. Concluded difficult to define single data-model combination for runoff simulation in West Africa, including Niger Basin. Also found for study area neither model fully represented basin characteristics

3. METHODS

In the following section, the evaluation methodology for the hydrological modelling framework, DECIPHeR (Dynamic fluxEs and ConnectIvity for Predictions of HydRology, Coxon et al., 2018) application to large and data sparse regional domains, using the Upper Niger and Inner Niger Delta as a case study, is described. Firstly the study location is described. This is then followed by a description of the modelling framework, and the experimental design for this study. Lastly, the input data used to force the hydrological model are described and discussed.

3.1. HYDROLOGY OF THE NIGER RIVER BASIN

The Niger River basin is located in Western Africa. Approximately 4200km in length, it is the largest river in western Africa, and the third largest on the African continent (Olomoda, 2002; Pedinotti et al., 2012), and has more than 100 million inhabitants within the 2.27 million km^2 basin area (World Bank, 2005). The basin spans across ten countries, with the largest sections in Mali, Niger and Nigeria, each contributing approximately 25% to the total basin area (Olomoda, 2002; Zwarts et al., 2005). The area of the basin in Guinea and Ivory Coast together only make up 5.3% of the total catchment area, however, as these are the locations of the sources and headwaters of the Niger River, they play a crucial role in defining the hydrology of the basin. The Inner Niger Delta is a key feature of the Niger Basin. This is a large (one of the largest in Western Africa) flooplain wetland, which can have an inundation area of approximately 40,000km^2 (Zwarts, 2010). The volume of water that enters Mali downstream of Guinea is actually greater than that enters Nigeria from Niger, approximately 1800km downstream (Zwarts et al.. 2005). This reduction in river flow is mainly due to the decline in runoff in the Inner Niger Delta, caused by large amounts of evaporation and the lack of runoff from the left bank of the river in Mali and Niger, as this is in the Sahara desert (Zwarts et al., 2005).

The Niger River basin has four main sub-regions, which are defined by their very different hydrological and/or climatic characteristics (Aich et al., 2016). This study focuses on the Upper Niger and Inner Niger Delta, as this stretch of the river has the most dynamic and variable hydro-climatic regime. Figure 3.1 shows the Upper Niger basin's location, the gauging stations in the catchment, the location of the Inner Niger Delta, and the major dams that were in operation during the simulation time period of this study.

<u>CLIMATE VARIABILITY OF THE UPPER NIGER BASIN</u>

The Inter-Tropical Convergence Zone controls the climate of this region, and therefore defines the hydrological characteristics of the rivers (Zwarts, 2011; Thompson et al., 2017). Mean annual rainfall across the Upper Niger Basin and Inner Niger Delta varies from approximately 2100 mm in the headwaters in Guinea, 1500mm in the Upper Bani (one of the major tributaries of the Upper Niger basin), to around 250mm in the downstream areas of the basin of the Inner Niger Delta (Thompson et al., 2016). Precipitation in the Upper Niger Basin is extremely seasonal; precipitation peaks in August across the region, but the annual wet season varies from 8 months (March-October) in the headwaters, to 3 months (July-September) in the Inner Niger Delta. The annual dry period is characterised by very little or no rainfall (Liersch et al., 2013). The inter-annual variability in precipitation is also large, but with a decline being reported since the 1970s (Zwarts et al., 2005; Mahe et al., 2009; Louvet et al., 2011). Spatial variation in potential evapotranspiration (PET) has a similar spatial variation, but the Inner Niger Delta experiences higher volumes of evaporation compared to the upstream headwaters. Thompson et al. (2016) report these volumes as 1800mm/year in the upstream regions, increasing to around 2150mm/year in the Delta area (PET was based on the Hargreaves method (Hargreaves and Samani, 1982) and using CRU TS 3.0 data (Mitchell and Jones, 2005).

RIVER FLOW VARIABILITY OF THE UPPER NIGER BASIN

River flows in the Upper Niger basin are also very seasonal, with large amounts of inter-annual variability, much like precipitation. There is a declining trend in annual discharge volumes and is reported to likely be caused by a combination of the reduction of baseflow in the basin, due to declining annual precipitation (Mahe, 2009), and changes in land use in the area (Aich et al., 2016). Water levels in the basin's rivers begin to rise with the start of the annual wet season, and discharges upstream of the Inner Niger Delta peak in September (Zwarts et al., 2005). The annual flood can take 3-4 months to pass through the Inner Niger Delta, therefore discharge peaks downstream in November or December (Sutcliffe and Parks, 1989, John et al., 1993). River discharge is substantially reduced by evaporation from the Inner Niger Delta, which is a seasonally inundated wetland, formed of an extensive network of lakes, streams and swamps (Zwarts et al., 2005; Liersch et al., 2013). Mahe et al. (2009) reports that the mean annual loss of river flow in this area is approximately 40%. These losses, as expected, are much larger during wetter years where there is extensive inundation, and smaller during dry periods (Zwarts and Grigoras, 2005; Mahe et al., 2009).

INFLUENCE OF GROUNDWATER IN THE UPPER NIGER BASIN

Groundwater can often have a strong influence on the hydrology of river basins. However, there is variability in the evidence of the degree of control that terrestrial water storage has in the Upper Niger basin (Werth et al., 2017). In the headwaters of the basin, groundwater aquifers are shallow and local with very little recharge potential, whereas in the Inner Niger Delta region, there are significant aquifers present (Further information on global groundwater aquifers are collected by the World-wide Hydrogeological Mapping and Assessment Program; WHYMAP, www.whymap.org). To investigate the role that aquifers have in defining the groundwater storage in the Inner Niger Delta, a comparison between the stable isotope composition of river water and groundwater was made (Fontes et al., 1999, p.199). It was demonstrated what while groundwater recharge may have occurred during the extensive flooding during wetter periods that characterised the Holocene, the recharge rates of today's climate are extremely low relative to the rainfall and evaporation in the region (Fontes et al., 1999; Dadson et al., 2010).

ANTHROPOGENIC ACTIVITIES AND INFLUENCE IN UPPER NIGER BASIN

There is also a large anthropogenic influence on the river flows of the Upper Niger. A number of water management schemes have been implemented in this region. There are three dams in the Upper Niger basin, which partially regulate the flow. The Sotuba dam has been in operation since 1929. It is a small hydropower dam, just downstream of Bamako. However, it has only a small reservoir, and its impacts on the hydrological characteristics of the Upper Niger are limited (Zwarts et al., 2005). The much larger Sélingué dam has been in operation since 1982 and is located on the Sankarani tributary. This dam has much larger effects on the hydrological regime of the Upper Niger, as Zwarts et al. (2005) have reported that it reduces peak flows at Koulikoro by 10-20% and 20-30% in wet and dry years, respectively. The dry season releases are significant and intend to maintain the downstream flows. The Markala Barrage diverts river flow from the main channel for irrigation on the Office du Niger project. Zwarts et al. (2005) suggest that these diversions are a small percentage of flows in the annual wet season but can reduce river flows by approximately half in the dry season.

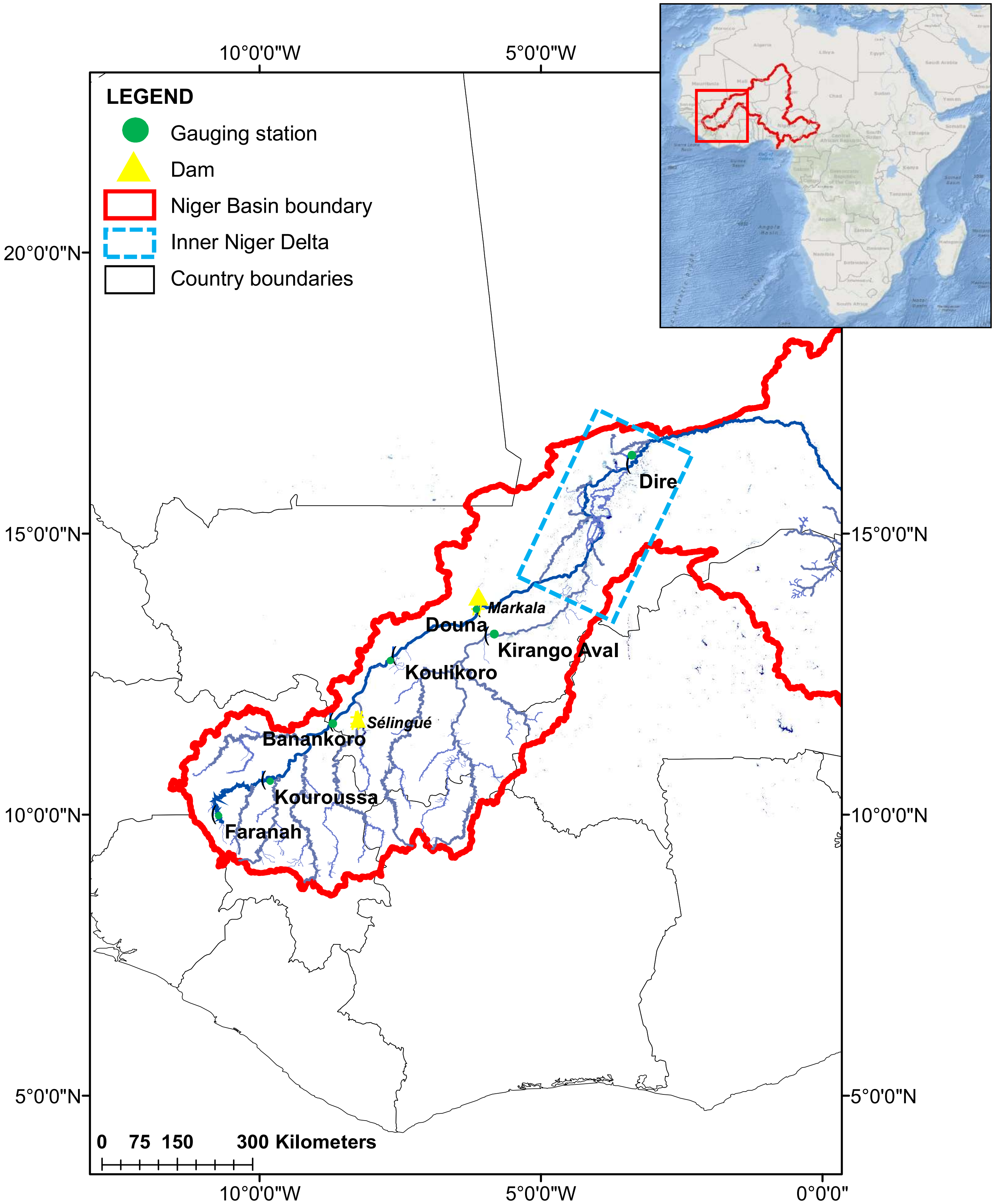


Figure 3.1. Map of the Upper Niger basin, including the gauging stations and major dams in operation in the study time period. The Inner Niger Delta is also highlighted.

3.2. EXPERIMENTAL SETUP

3.2.1. DECIPHeR

DECIPHeR (Dynamic fluxEs and ConnectIvity for Predictions of HydRology, Coxon et al., 2018) is a flexible hydrological modelling framework for uncertain flow simulation and prediction at catchment to continental scales. It builds on the key concepts of Dynamic TOPMODEL, originally developed by Beven and Freer (2001a). Dynamic TOPMODEL has mostly only been applied to single catchments. Previously, hydrological modelling studies were constrained by a lack of hydrological datasets and computing power. However, recent advancements in the availability of global open-source datasets and high-performance computing has allowed for the development of flexible hydrological modelling frameworks that can represent small scale features of catchments and can produce large ensembles of discharge predictions at a range of scales. DECIPHeR can be applied from catchment to continental scale and is a flexible, computationally efficient modelling framework that can be adapted for specific hydrologic settings or data availability.

The DECIPHeR modelling framework has so far only been demonstrated for use in large scale applications in the initial model evaluation in the UK – a data rich and well gauged location. Therefore, to evaluate model performance in large and data scarce river basins, DECIPHeR is applied to the Upper Niger Basin in this study. This is a suitable location for investigating model performance and for highlighting areas for future model improvement and development as the data available in this basin is from global open-source datasets and there are highly varying hydroclimatic conditions from upstream to downstream, as described in Section 3.1.

There are a number of key features of DECIPHeR which make it a suitable hydrological modelling framework for the Upper Niger Basin. Firstly, the model build can be fully automated, which allows it to be applied across large scales easily. Secondly, DECIPHeR is a flexible modelling framework, which allows for experimentation and hypothesis testing of different basin functioning by using different models structures and parameterisations to represent this. This flexibility is an especially valuable attribute in river basins, such as the Upper Niger, where there are few or no ground observations and measurements to inform us about catchment functioning and processes. DECIPHeR is also computationally efficient and has the capacity to run large ensembles for uncertainty analysis and hypothesis testing. Finally, DECIPHeR uses hydrological response units (HRUs) to group together regions in the study location that have similar characteristics and attributes. These parts of the landscape maintain hydrological connectivity and can be mapped back into space.

HRUs are used to discretise the model domain in the DECIPHeR framework, instead of using gridded or fully distributed approaches, which are much more computationally demanding. The user can choose to split the landscape up in any configuration, with any level of spatial complexity, that they wish/see fit for purpose. To define similarity within the landscape, user supplied data is used, for example different catchment characteristics, such as topography, geology, land use, soils, and/or spatially varying inputs, such as rainfall and evapotranspiration. This means that areas across the catchment that share the same landscape attributes and/or climatic variability will be grouped together, and treated homogenously in the rainfall runoff modelling. This minimises model simulation times, and allows for this flexible modelling framework's computational efficiency.

There are two key stages in the DECIPHeR framework: 1) A digital terrain analysis (DTA) is carried out to define the gauge network, river network and river routing, define the HRUs and discretise the model domain, and to specify the spatial heterogeneity and hydrological connectivity within the catchment, and 2) this model domain is then run in the rainfall-runoff model to simulate flow timeseries at the user specified locations. These two stages are described in the following sections and how they were implemented for the Upper Niger basin.

3.2.2. DIGITAL TERRAIN ANALYSIS

In DECIPHeR, the model domain is built in the DTA. This is critical as it defines the HRUs, characterises hydrological connectivity in the landscape, and to set up the river network and routing scheme from which flow timeseries will be outputted. The minimum data requires for the DTA is a digital elevation model (DEM) and XY locations of where simulated discharge is required during the rainfall-runoff stage. These locations can be any gauged or ungauged points on the river network. There is the option for additional data to be incorporated in more data rich areas and depending on the objectives of the modelling study. These include geology, land use, land cover, soil types, and climatic variables. Figure 3.2 summaries the key steps in the DTA with examples of the outputs for the Upper Niger basin.

DIGITAL TERRAIN ANALYSIS FOR THE UPPER NIGER BASIN

To run the DTA for the Upper Niger basin, the global gridded MERIT DEM (Yamazaki et al., 2017) was used. This is the most accurate DEM that is available for the region, at approximately 90m resolution at the equator. It was developed by removing major errors, including absolute bias, stripe noise, speckle noise and tree height bias using multiple satellite datasets and filtering techniques, from existing DEMs, NASA SRTM3 DEM (Farr et al., 2007) and the JAXA AW3D (Tadono et al., 2015). It was found that previous DEMs were misrepresenting the topography in many of the world's major river basins, including the Niger (Yamazaki et al., 2017), and therefore this new DEM has improved the accuracy in these regions. This DEM was also pit and sink filled in order to remove any flat areas and maintain important topographic features before running in the DTA.

A river network for the Upper Niger basin was built in the DTA that matches the flow direction of the DEM. This is generated from a list of headwater cells in the model domain that are then routed downstream via the steepest gradient, where all rivers will connect to the boundary of the DEM. As there was no reference river network from which these headwater cells could be derived for the Upper Niger basin, these headwater cells were identified from the DEM by calculating topographic index and accumulated area for the whole of the model domain. Thresholds in these values where applied (i.e. cells that have values higher than the threshold in both variables is classified as a river cell) in order to generate the river map. A trial and error approach to choosing these threshold values was used. The river map generated with different thresholds was inspected each time, and the values that produced the most reasonable river network were chosen. These values were deemed to be most optimal for thresholds of 15 for topographic index, and 1200000 for accumulated area.

The next stage was to identify the locations on this generated river network where simulated flow timeseries is to be outputted. The locations used were the XY coordinates for the gauging stations on the Upper Niger basin that had the longest daily discharge records, with observation periods that overlapped with each other. In total, there are seven gauging stations that met these criteria, and the location of each of these gauges can be seen in Figure 3.1. The best candidate cell in the river network was chosen for each of these gauges by using a search radius of 2km. This search radius was set to be wider than the default of 500m in the DTA code as this river basin is very large. The closest river cell was chosen for each gauge. Catchment masks were then created for each of these gauges. All river cells are then linked together to build the routing tree, and this forms the basis of the river network that is used in the rainfall-runoff modelling stage.

The final stage of the DTA is to classify the HRUs in order to discretise the model domain. For the Upper Niger basin, topographic information and spatially varying inputs were used to define hydrological similarity across the landscape. This is due to the lack of available data in the region for different catchment characteristics, such as geology, soil types, land use etc, and also to maintain computational efficiency in such a large river basin. A 0.5°x0.5° rainfall and PET grid was used to represent the climatic variability across the Upper Niger basin, and three equal percentiles of slope and accumulated area, as calculated from the DEM, were used to define similarity within the basin. In total, 1843 HRUs were classified from the input data. Figure 3.3 shows how the HRUs were determined for the Upper Niger basin.

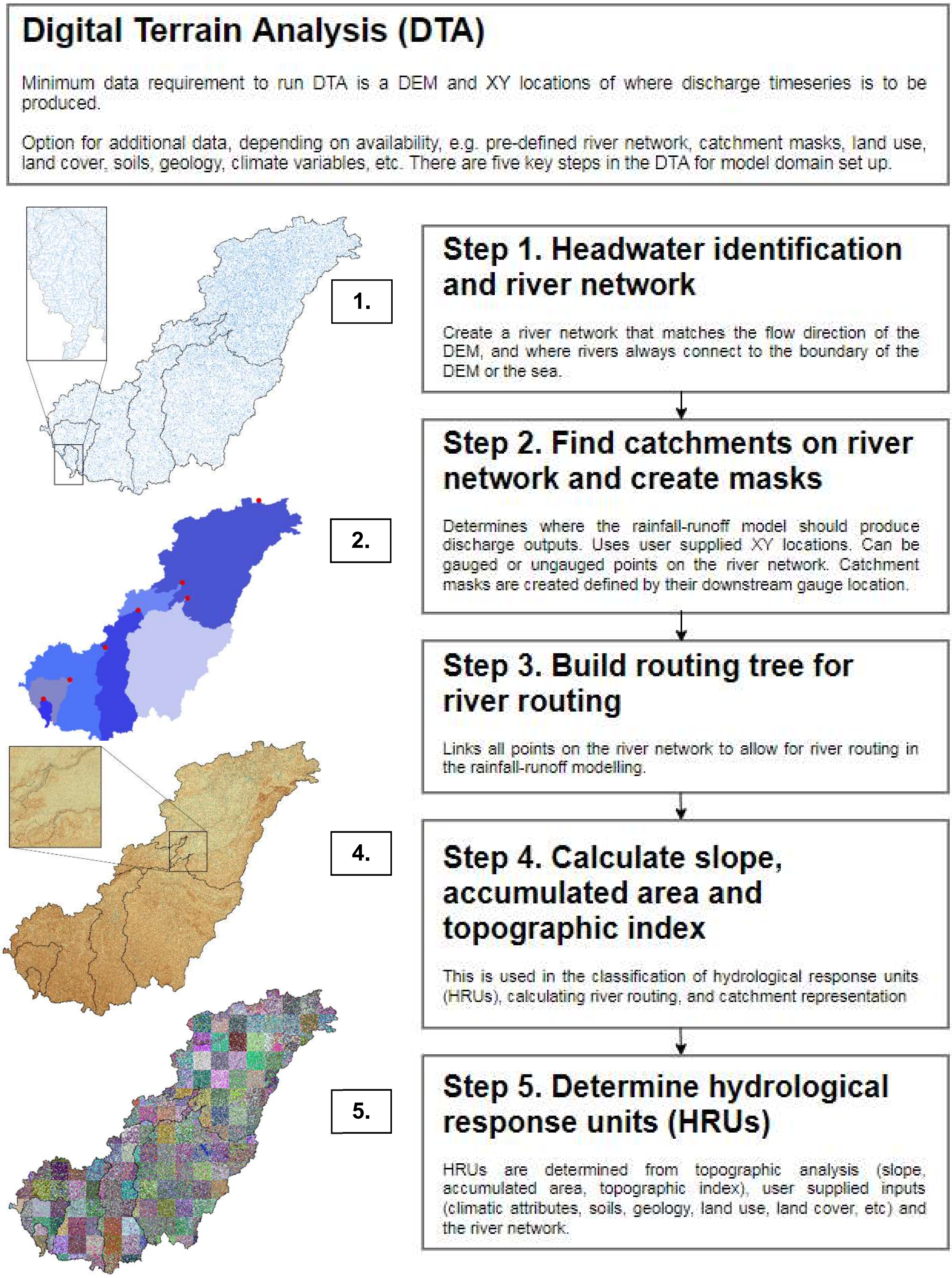


Figure 3.2. Processing steps for the Digital Terrain Analysis (DTA). Examples of outputs at each step for the Upper Niger basin are given.

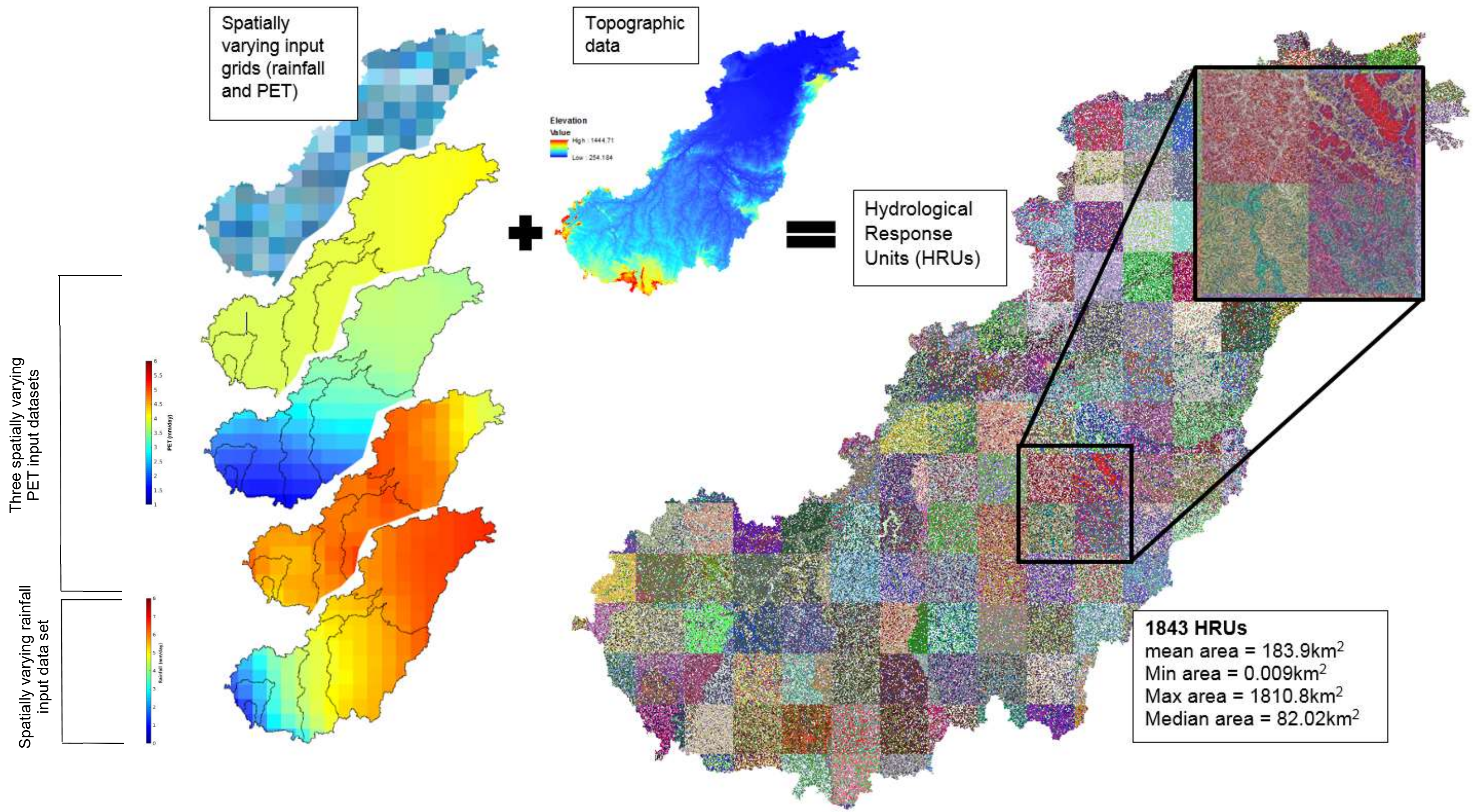


Figure 3.3. Digital Terrain Analysis for the Upper Niger basin. Hydrological response units are determined by topographic data, gridded precipitation and PET data, and three equal fractions of accumulated area and topographic index.

3.2.3. RAINFALL RUNOFF MODELLING

On completion of the DTA, the second key stage in the DECIPHeR framework is rainfall-runoff modelling. HRUs that have been classified in the DTA are run in the hydrological model to produce simulated flow timeseries at the chosen points in the model domain. The model parameters and model structure will be described in the following sections.

MODEL PARAMETERS

DECIPHeR can be run using either a default value for each parameter across all HRUs for all timesteps or using Monte-Carlo sampling of a parameter space between user defined upper and lower boundaries. In the DTA, the user is given the option for how these parameters will be implemented in the rainfall-runoff model simulations. For the simplest case, parameters are applied homogenously across the model domain. There is also the option to experiment with different parameter values in different parts of the landscape. There are seven parameters currently used in the model. These parameters are given, with a description of their function, in Table 3.3.

Table 3.3. Seven parameters are used in the hydrological model. Parameter names, units and a description are given.

PARAMETER	DESCRIPTION
SZM [m]	Form of the exponential decline in conductivity
Ln(T_0) [m^2 h^{-1}]	Effective lateral saturated transmissivity
SRmax [m]	Maximum root zone storage
SRinit [m]	Initial root zone deficit
CHV [m h^{-1}]	Channel routing velocity
T_d [m h^{-1}]	Unsaturated zone time delay
Smax [m]	Maximum effective deficit of subsurface saturated zone

MODEL STRUCTURE

The current DECIPHeR model formulation implements a single model structure for each HRU. Figure 3.4 shows a conceptual diagram of the initial model structure, described in Coxon et al. (2018), which is implemented for each individual HRU. The model parameters are summarised in Table 3.3, and the stores and fluxes are summarised in Table 3.4.

The model structure has three stores: the soil root zone (S_{RZ}), the unsaturated zone (S_{UZ}), and the saturated zone. The hydrological model code loops through all HRUs, starting off with the soil root zone. Precipitation is added directly to the soil root zone. If the soil root zone is at field capacity (i.e. it is full, and this is determined by the value of parameter SRmax, which defines what the maximum storage in the soil root zone can be), then the excess is transferred to the unsaturated zone. Actual evapotranspiration (AET) is calculated for each HRU and is taken directly from the soil root zone, and this will vary across the basin depending on the spatially varying input PET data. It is calculated as a proportion of the water storage in the soil root zone. If the soil root zone is at its maximum capacity, then AET is taken at the full potential rate, as defined by the input PET for each timestep. If, for example, the soil root zone is 50% full, then AET is taken as 50% of the input PET.

Once the soil root zone has reached its maximum holding capacity, any excess rainfall is routed to the subsurface (S_{UZ}). If this store is full, then the excess is added to the saturation excess storage (S_{EX}). This store is a combination of the excess of two internal fluxes – the saturated excess flow (Q_{EXS}) and the precipitation excess flow (Q_{EXUS}). Q_{EXS} represents infiltration excess overland flow (or Hortonian flow), where the rate of precipitation is higher than the rate of infiltration through the soil. Q_{EXUS} represents saturation excess overland flow, where the soil is saturated and any further precipitation immediately produces surface runoff. The saturation excess storage (S_{EX}) becomes an internal flux, overland flow (Q_{OF}), and this is then routed through the sub-basin to the outlet. If the water remains in the unsaturated zone, then it will be routed further through the subsurface (Q_{UZ}). Recharge from the unsaturated zone to the water table is at a rate proportional to the ratio of unsaturated zone storage (S_{UZ}) to storage deficit (S_D) and the gravity drainage time delay parameter, T_d.

The saturated zone is controlled by the changes in storage deficits (S_D). The storage deficit for each HRU is calculated from the drainage from the unsaturated zone (Q_{UZ}), the inputs from upslope HRUs (Q_{IN}) and downslope flows out of each HRU (Q_{SAT}). The distribution of fluxes between HRUs in the subsurface is defined by the flow weightings matrix derived in the DTA. Transfer of storage between HRUs is calculated using a kinematic wave approximation. This is controlled by the parameter SZM, which forms the exponential decline in hydraulic conductivity in the saturated zone. This in turn affects the shape of the recession curve and $\ln(T_0)$ which controls the lateral transmissivity. If maximum deficit of the subsurface saturated zone is reached, (i.e. the storage deficit is greater than Smax) then there is no downslope flow between HRUs. If the storage deficit is less than zero, then any water added to the storage is routed as saturation excess overland flow.

The model produces simulated discharge for each river ID at each timestep. River IDs are determined in the creation of the river network during the DTA. A time delay needs to be added to this flow as water entering the river channel upstream will exit at the downstream outlet at a slower rate than any water that enters further downstream. A fixed channel wave velocity (CHV) is applied throughout the network to account for delay and attenuation. This parameter is found to be particularly sensitive in large catchments as it aids in the timing of events.

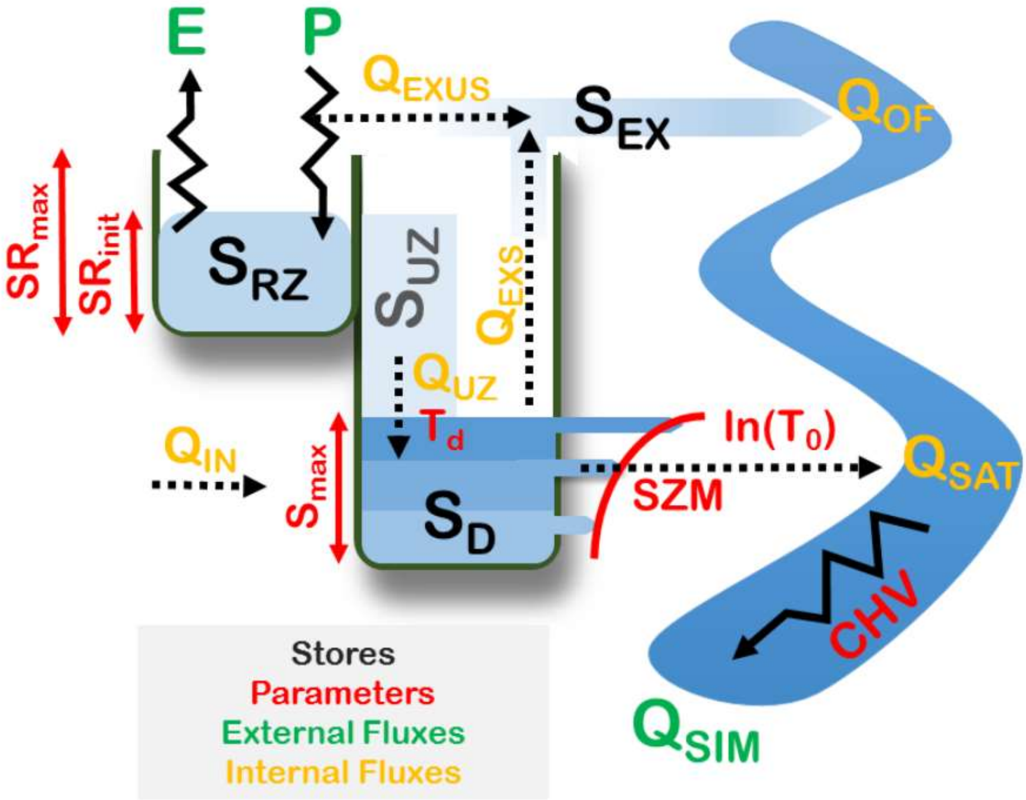


Figure 3.4. Conceptual diagram of the current model structure implemented in DECIPHeR, taken from Coxon et al. (2018). The modelling framework allows for flexibility in this structure, and therefore can easily be modified by the user for different conceptualisations of the hydrological processes and dynamics. All parameters are described in Table 3.3. and stores and fluxes are described in Table 3.4.

Table 3.4. Stores, internal and external fluxes in the DECIPHeR models structure.

STORES	S_{RZ}	Root zone storage	m
	S_{UZ}	Unsaturated storage	m
	S_{EX}	Saturation excess storage	m
	SD	Saturated storage deficit	m
INTERNAL FLUXES	Q_{UZ}	Drainage flux	m ts^{-1}
	Q_{IN}	Upslope input flow	m ts^{-1}
	Q_{EXS}	Saturated excess flow	m ts^{-1}
	Q_{EXUS}	Precipitation excess flow	m ts^{-1}
	Q_{OF}	Overland flow (Q_{EXS} + Q_{EXUS})	m ts^{-1}
	Q_{SAT}	Saturated flow	m ts^{-1}
EXTERNAL FLUXES	P	Precipitation	m ts^{-1}
	E	Potential evapotranspiration	m ts^{-1}
	Q_{OBS}	Observed discharge	m ts^{-1}
	Q_{SIM}	Simulated discharge	m ts^{-1}

3.2.4. EXPERIMENTAL DESIGN

To initially evaluate DECIPHeR in the Upper Niger basin, the model structure that is used in Coxon et al. (2018) was applied to all HRUs in the model domain as a benchmark of model performance. Initially, one gridded daily precipitation input dataset was used (MSWEP, Beck et al., 2018), and one gridded daily PET dataset was used (GLEAM, Martens et al., 2017), both with spatial resolution of 0.5°x0.5°. The model was run within a Monte-Carlo simulation framework, where 10,000 parameter sets were randomly sampled from an assumed uniform distribution. These parameter sets were implemented with spatial uniformity across the whole model domain, i.e. each HRU has the same parameter values in all sub-basins. This number of parameter sets was chosen as it provides a representative sample of the parameter space and a reasonable model run time (approximately 6 days for 10,000 parameter sets). However, it should be noted that this is a limiting factor in the evaluation of the model. A larger sample of the parameter space would reduce parametric uncertainty, but this would largely decrease the computational efficiency of the model in such a large river basin. The upper and lower boundaries from which the parameters are sampled from are given in Table 3.5 which are the same as used in Coxon et al. (2018), as these limits were chosen to represent the highly varying hydroclimatic conditions in a national application of DECIPHeR, and are therefore set to be wide to allow for a wide parameter space. For each model simulation, the parameter sets were applied homogeneously across all HRUs in the basin and used with a single homogenous model structure (which is described in Section 3.2.3).

In previous hydrological modelling studies in the Niger basin, it has been found that model performance is sensitive to an underestimation or misrepresentation of evaporation (e.g. Schuol and Abbaspour, 2006; Dadson et al., 2010; Anderrsson et al., 2017a). Therefore, two additional gridded daily PET input datasets were also used to force the model – ECMWF EartH2Observe (Schellekens et al., 2017), and a constructed PET grid based on a simple relationship between daily mean temperature and latitude (Oudin et al., 2005), both also with 0.5°x0.5° spatial resolution. Again, to explore this representation of evaporation in the model simulations, a second iteration of model simulations were produced, where the upper bound for parameter SRmax was increased, to create a larger soil root zone. It is hypothesised that this will allow for more AET to be taken from this store. This model set-up was run using the same daily precipitation and three daily PET grids, for 10,000 randomly sampled parameter sets.

Next, a modified model structure was implemented, which incorporates evaporative losses from HRUs from the saturated zone to allow for a higher proportion of AET to be taken from HRUs. Figure 3.5 shows a conceptual diagram of this new model structure. In this modified structure, evaporative demands are initially satisfied by the soil root zone, and any residual evaporative demand can then be satisfied by the saturated zone. Firstly, the residual evaporative demand is calculated

$$resid\_evap = PET - ET$$

Where ET is evaporation from the soil root zone. Next, the available storage in the saturated zone is calculated

$$satstore = Smax - SD$$

In order for residual evaporative demands to be taken from the saturated zone, resid_evap and satstore must be greater than 0. If this condition is met, satevap is calculated in a similar way at evap from the soil root zone, where ET is proportional to how full the store is

$$satevap = resid\_evap * (satstore / Smax)$$

Or if resid_evap = satstore,

$$satevap = satstore$$

The saturated deficits for each time step are updated and the amount of evaporation at each timestep is also updated

$$SD = SD + satevap$$

$$ETout = ET + satevap$$

This evaporation is then accounted for as an additional negative flux in the kinematic wave formulation, to account for the losses from the saturated zone in the transfers of water to downslope HRUs. This model structure was run with each of the PET products with the same 10,000 parameter sets that were used in the initial model simulations, and then the same 10,000 parameter sets that were used in the model simulations with the increased SRmax upper bound.

Table 3.5. DECIPHeR parameter ranges for initial application to the Upper Niger basin.

PARAMETER	UNITS	LOWER BOUND	UPPER BOUND
SZM	m	0.001	0.07
SRmax	m	0.005	0.15
SRinit	m	0	0.01
T_d	m hr^{-1}	0.1	40
CHV	m hr^{-1}	250	4000
$Ln(T_0)$	$Ln(m^2\ hr^{-1})$	-7	5
Smax	m	0.2	3

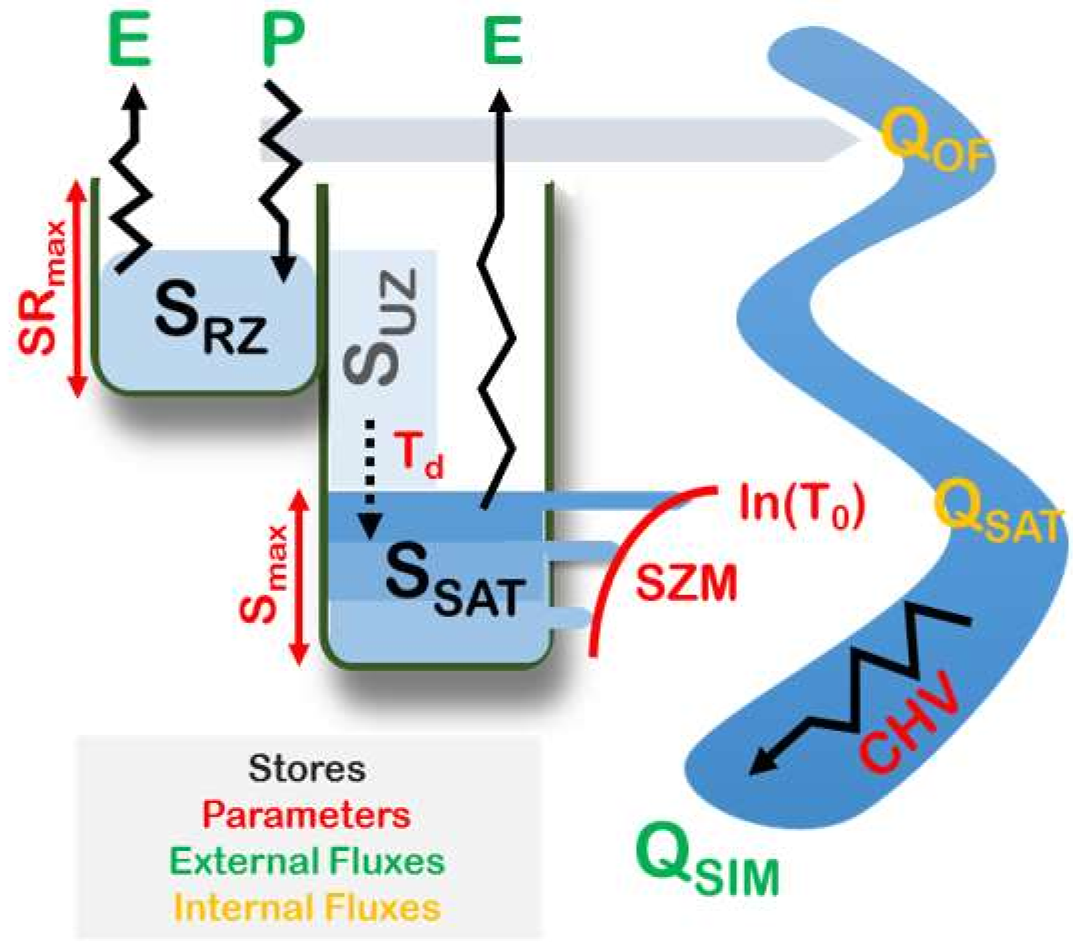


Figure 3.5. Conceptual diagram of the modified model structure to include evaporation from the saturated zone, adapted from Coxon et al. (2018).

3.3. INPUT DATA

The input data requirements for DECIPHeR and the input datasets that were used in this study are mentioned in Section 3.2. These datasets will be discussed in more detail in this section.

The data required to force the rainfall-runoff model are a timeseries of precipitation and PET, and discharge observations for the initialisation of the model. In this study, data from global open-source datasets was used. A simulation time period of 21 years (01/01/1980 – 31/12/2000) was chosen. This period covers the dry, drought years that characterised the 1970-80s (Paturel et al., 2007) and the period that followed where rainfall in the region began to increase again (Zwarts, 2010). A summary of the input data for this application of DECIPHeR to the Upper Niger basin are summarised in Table 3.6, with a description of the data manipulation that was required in order for it to be used in the model.

Table 3.6. Data sources and manipulations (pre-processing) made to all data. All hydrological data had to be converted to m/day for use in DECIPHeR.

DATA FOR DECIPHeR	UNITS	ADDITIONAL INFO	MANIPULATIONS	SOURCE
Topographic data	m	Terrain elevations at a 3sec resolution (~90m at the equator)	The relevant 5°x5° tiles were mosaicked together in ArcMap. Converted to UTM zone 28N coordinates	MERIT DEM (Yamazaki et al., 2017)
Daily precipitation	mm/day	Gridded, 0.1° spatial resolution, 1979-2016	Aggregated to spatial resolution of 0.5°, converted to m/day	MSWEP (Beck et al., 2018)
Daily potential evapotranspiration	mm/day	Gridded, 0.25° spatial resolution, 1980-2017	Aggregated to spatial resolution of 0.5°, converted to m/day	GLEAM (Martens et al., 2017)
	Kg m^2 s^{-1}	As above	As above	ECMWF EartH2Observe (Schellekens et al., 2017)
	mm/day	Daily PET estimate based on simple relationship between daily mean temperature and latitude	Created gridded PET with resolution 0.5°, from gridded daily mean temperature and latitude of the grid cells	(Oudin et al., 2005) WATCH forcing data (Weedon et al., 2014)
Daily discharge	m^3/s	Daily mean discharge, varying start and end dates for gauges across basin	Converted to m/day	Global Runoff Data Centre (GRDC)

3.3.2. DISCHARGE DATA

Discharge data was taken from the Global Runoff Data Centre (GRDC, 2018), which operates under the authority of the World Meteorological Organisation. This is an international archive of river discharge data up to 200 years old, and has allowed for multi-national and global, long-term hydrological studies. The aim when the database was created was to facilitate earth scientist's analysis of global climate trends and assess environmental impacts and risks. Daily and monthly discharge data is collected for around 9,500 stations in 160 countries, with an average record length of 43 years (GRDC, 2018). However, the quality of the discharge data in Western Africa in this database has been declining since the 1980s when many hydrometric stations were closed and reduced to a minimum (Nkamdjou and Bedimo, 2008).

There are only seven gauging stations in the Upper Niger basin with near complete daily discharge data for the simulation period 01/01/1980 – 31/12/2000. Table 3.7 gives the names and station codes for the gauges and some of the catchment characteristics of the sub-basin defined by their downstream station.Table 3.8 summarises the availability and completeness of discharge records for the seven gauging stations in the study location.

Table 3.7. Summary information, in upstream to downstream order, about the gauging stations in the Upper Niger basin, taken from GRDC (2018)

STATION NAME	COUNTRY	LATITUDE	LONGITUDE	ELEVATION (m)	BASIN AREA (km^2)	AVALIABLE DAILY DATA
Faranah 1634200	Guinea	10.33	-10.75	456	3180	1955-2001
Kouroussa 1634400	Guinea	10.65	-9.87	362	14,820	1923-2002
Banankoro 1134030	Mali	11.68	-8.66	329	53,750	1967-2001
Koulikoro 1134100	Mali	12.87	-7.55	290	48,250	1907-2006
Douna 1134300	Mali	13.22	-5.90	270	101,600	1922-2001
Kirango Aval 1134250	Mali	13.72	-6.05	274	17,000	1925-2001
Dire 1314700	Mali	16.27	-3.38	256	101,400	1924-2003

Table 3.8. Available daily discharge data for gauging stations in Upper Niger basin, and showing the completeness of the record for each year of simulation.

STATION	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Faranah 1634200	<10%	<10%	<10%	<10%	<10%	<10%	>50%	>50%	<10%	<20%	<10%	<10%	<10%	>50%	<10%	<10%	<20%	<10%	<10%	<50%	<10%
Kouroussa 1634400	<10%	<10%	<10%	<10%	<10%	<20%	<10%	<10%	<20%	<10%	<10%	<20%	<50%	<50%	<50%	<50%	<50%	<10%	<10%	>50%	<50%
Banankoro 1134030	<50%	<50%	<50%	>50%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<50%	<10%	<10%	<10%	<50%	>50%	<10%	<10%
Koulikoro 1134100	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%
Douna 1134300	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	>50%	<10%	<10%	<10%	<10%	<10%	<20%	<50%
Kirango Aval 1134250	<50%	<10%	<10%	<20%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	>50%	<10%	<50%	<20%	>50%	>50%	>50%	>50%
Dire 1134700	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	<10%	>50%	<10%	<10%	<10%	<50%	<10%	<50%	<50%

<10% missing	**<20% missing**	**<50% missing**	**>50% missing**

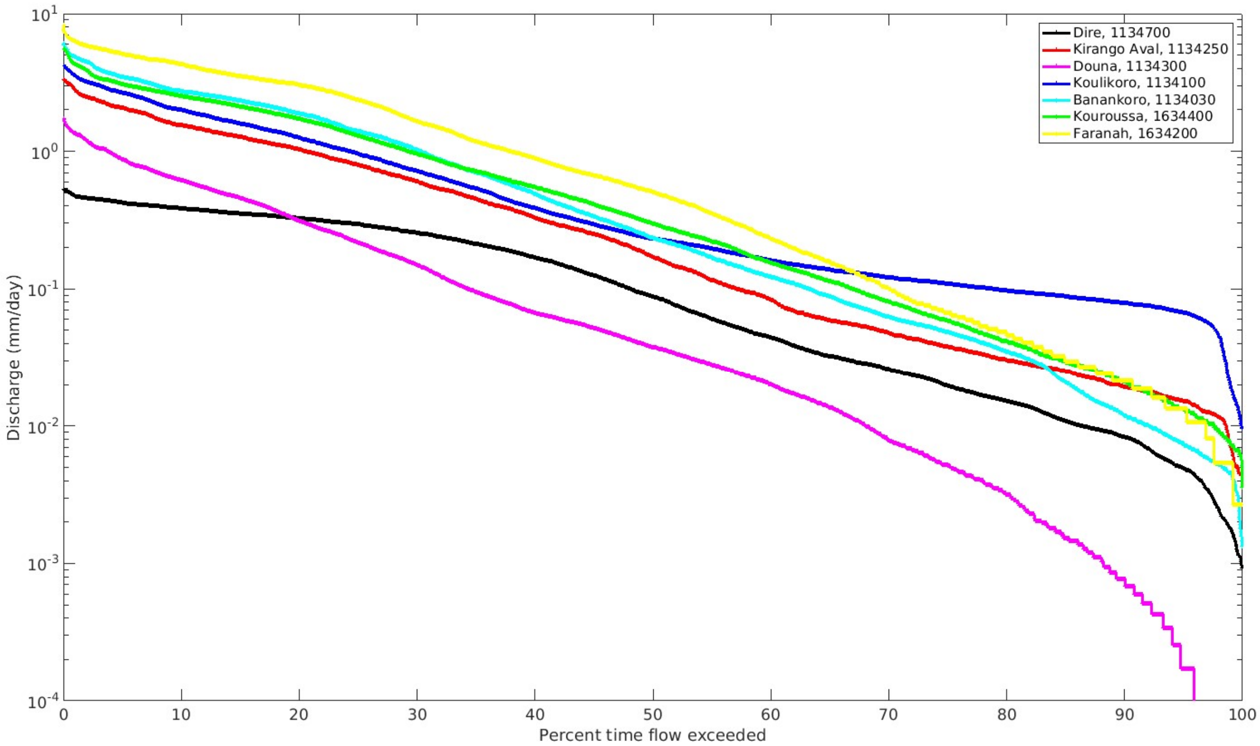


Figure 3.6. Flow duration curves for each of the sub-basins, defined by their downstream gauging station, taken from the discharge time series.

3.3.3. PRECIPITATION DATA

Precipitation is one of the key drivers of the hydrological cycle, and is a critical aspect in any hydrological modelling study (Kidd and Huffman, 2011; Hou et al., 2014). Accurate representation of rainfall in models is needed to produce accurate simulations of streamflow (Beven, 2004). The rain gauge network in regions outside of developed countries is very sparse, and data records are of low quality or incomplete. The development in satellite and remote sensing based observations of precipitation has offered a solution to this issue. Global open-source datasets have been established, and this has provided hydrologists with the means of modelling in data scarce regions of the world. Table 3.9 summarises some of the global datasets that contain precipitation. This table shows that there are different data sources for these different global products, and a variety of temporal and spatial resolutions, covering a range of different time periods. For the Upper Niger basin, a precipitation product that has daily temporal coverage was needed for a more robust evaluation of the performance of DECIPHeR. Spatial resolution needs to be as fine as possible in order to capture the spatial variability of rainfall patterns across the basin. Multiple-Source Weighted-Ensemble Precipitation was chosen for use in this modelling study, and this dataset will be discussed in the following section.

Table 3.9. Summary of some of the global datasets that contain precipitation observations, with the source of the data and the spatial and temporal resolution.

DATA SET	DATA SOURCE	SPATIAL AND TEMPORAL RESOLUTION	REFERENCE
GPCC Global Precipitation Climatology Centre monthly precipitation dataset	Gauge	Monthly, 1901-2001 Global grids: 0.5º, 1.0º, 2.5º	Beck et al., 2005
Global Historical Climatology Network Daily Database	Station records	Daily, 1861-present. Records from approx. 80,000 stations in 180 countries.	Menne et al., 2012
GPCP (Daily): Global Precipitation Climatology Project-1DD product	Infrared satellite	Daily rainfall globally from October 1996. 1º global grid.	Pendergrass and Hartmann., 2014
TRMM - Tropical Rainfall Measuring Mission	Satellite	3 hourly from 1998. 0.25º girid 50ºS-50ºN.	Huffman et al., 2007, 2010
CRU TS3.10	Station records	Monthly 1901-2016. 0.5º global grid.	Harris et al., 2014
GPM – Global Precipitation Measurement Mission	Satelitte	Half-hourly precipitation estimates. 0.1º gridded data for 60ºN-60ºS	NASA, 2011
NCEP Climate Forecast System Reanalysis	Model reanalysis	6 hourly from 1979. 0.1º global grid	Saha et al. 2010
WFD – Watch Forcing Data	Based on ERA-40	3/6 hourly 1901-2001. 0.5º global grid	Weedon et al., 2011
WFDEI – WATCH Forcing Data methodology applied to ERA-Interim Data	Using ERA-Interim reanalysis data	Daily 1979-2012. 0.5º global grid	Weedon et al., 2014
PERSIANN-CDR – Precipitation Estimation From Remotely Sensed Information Using Artificial Neural Networks – Climate Data Record	Remote sensing	Daily from 1983. 0.25º grid 60ºS-60ºN	Ashuri et al., 2015
MSWEP Retrospective – Multi-Source Weighted-Ensemble Precipitation	Wide range of data sources: gauges, satellites, and model reanalysis	3-hourly or daily, 1979-2016 0.1º global grid	Beck et al., 2018

MULTI-SOURCE WEIGHTED-ENSEMBLE PRECIPITATION

Multi-Source Weighted-Ensemble Precipitation (MSWEP) is a new global precipitation data set for the period 1979-2015 that has been specifically designed for hydrological modelling (Beck et al., 2018), and was chosen as the rainfall input data for this project for this reason. It also has a higher spatial resolution than many other global precipitation data sets, at 0.1°x0.1°. MSWEP was designed to address the common issue current global precipitation datasets face, which is not taking advantage of satellite and reanalysis data that is available. The initial aim was to merge high quality precipitation data as a function of time and space. The long-term mean of MSWEP is based on Climate Hazards Group Precipitation Climatology (CHPclim) dataset (Funk et al., 2015), which is a global precipitation climatology based on gauging station and satellite data, but are replaced where more accurate regional datasets are available. In order to account for gauge under-catch and orographic effects, a correction was used by inferring catchment average precipitation for discharge measurements at 13,762 gauging stations across the world. The temporal variability was determined by calculating weighted averages of precipitation anomalies using seven datasets. Two of these are based on the interpolation of gauge measurements (CPC Unified and GPCC), three on satellite remote sensing (CMORPH, GSMaP-MVK, and TMPA 3B42RT), and two atmospheric model reanalyses (ERA-Interim and JRA-55). For every grid cell, weight assigned to the gauge-based estimates is calculates from the density of the gauging network in the region, and the weights assigned to the satellite- and reanalysis-based estimates are calculated from the surrounding gauges. In order to obtain optimal results, the weights of these different data sources varies spatially across the globe (Beck et al., 2017). To determine the quality of the MSWEP dataset, an independent precipitation data set from 125 FLUXNET tower stations (global network of micrometeorological towers) around the world were used for comparison and validation. MSWEP has also been applied in hydrological modelling studies in data sparse catchments (Alijanian et al., 2017; Nair and Indu, 2017; Shalou et al., 2017) where simulations have obtained good overall performance.

Figure 3.7 shows the spatial variability of mean daily rainfall estimates in MSWEP for the Upper Niger Basin for the simulation time period (1980-200), and Figure 3.8 shows the seasonal variability. This rainfall data shows a similar spatial pattern as is described in the literature (e.g. Zwarts et al., 2005; Zwarts, 2010; Tarhule et al., 2014 ; Thompson et al., 2016, 2017).

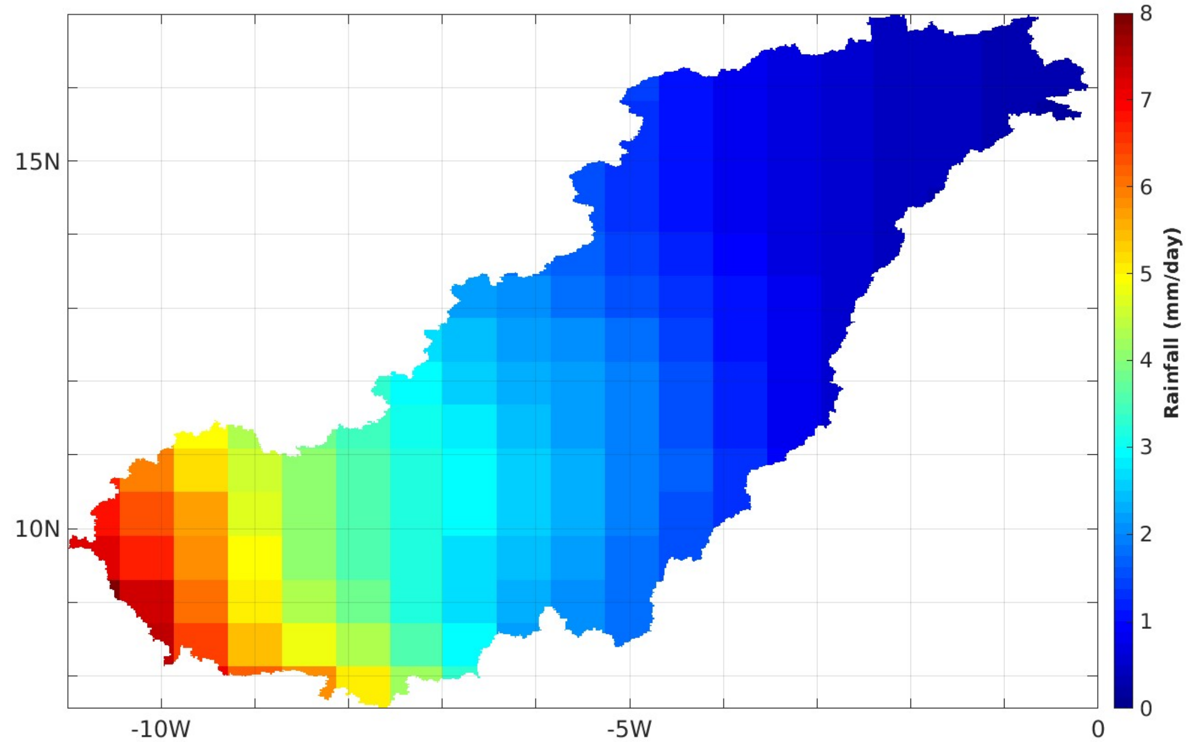


Figure 3.7. Spatial variability of MSWEP daily mean rainfall estimates (mm/day), 1980-2000.

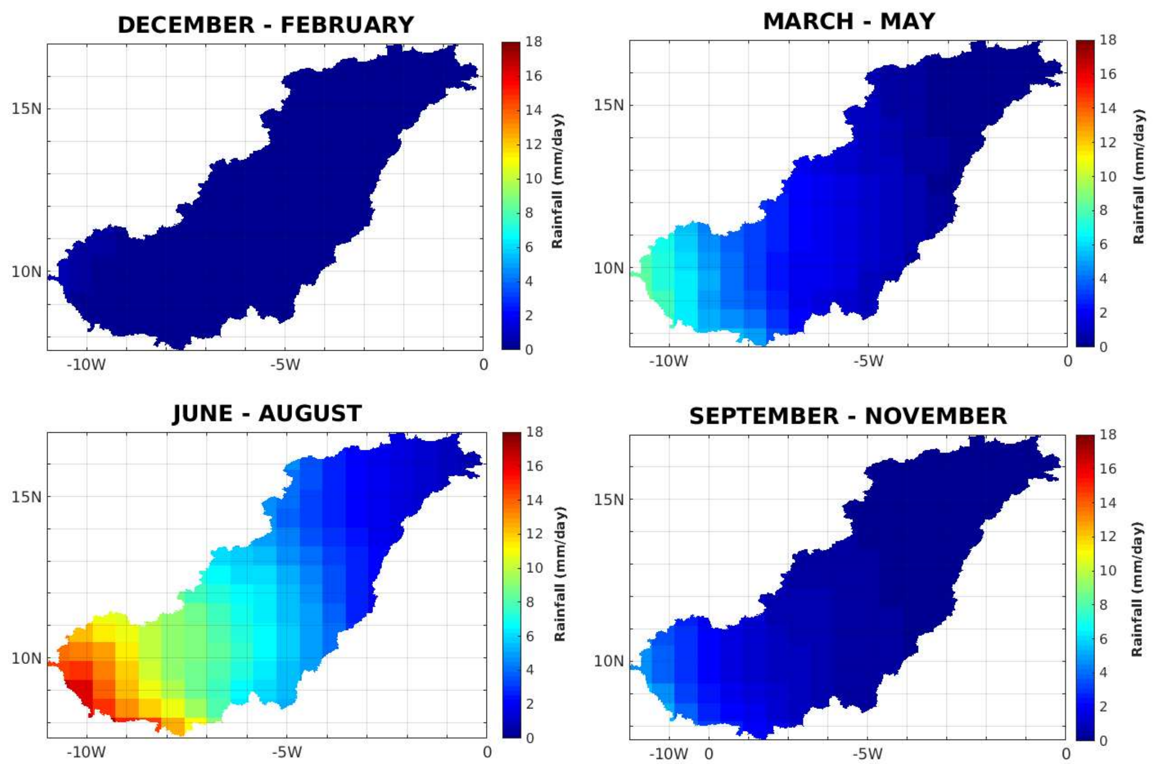


Figure 3.8. Seasonal variability of MSWEP daily mean rainfall estimates (mm/day), 1980-2000.

3.3.4. POTENTIAL EVAPOTRANSPIRATION DATA

Potential evapotranspiration (PET) is a major component of a catchment's water balance, and therefore this should be a key input to any hydrological model. However, it is common in rainfall-runoff model studies to use a lumped mean PET that has the same seasonally variability and is repeated each year (Oudin et al., 2005) instead of temporally varying PET. This is due to the lack of high quality meteorological data needed in most equations to estimate PET. Due to the hydroclimatic variability across the Upper Niger basin, PET is considered to be spatially and temporally variable, and therefore for this study, spatially varying gridded PET products were chosen. There are also large differences in the estimates of land-surface evaporation from current existing methodologies (Jimenez et al., 2011; Mueller et al., 2011) and this indicates that evaporative fluxes are one of the most uncertain elements of the global water cycle. To assess the uncertainty of different global PET products, two different datasets were used, as well as a constructed dataset based on a simple relationship between daily mean temperature and extra-terrestrial radiation. Figure 3.9 shows the mean monthly PET values, and Figure 3.10 shows the mean annual PET for all sub-basins in the model domain. From this figures, it is clear that there are large differences between the datasets, both in terms of their intra- and inter-annual variability. These datasets will be described below.

Table 3.10. Summary of some global evapotranspiration datasets, with the spatial and temporal resolution

DATA SET	METHOD	SPATIAL AND TEMPORAL RESOLUTION	REFERENCE
GLDAS – Global Land Data Assimilation System	Land surface model, water and energy balance. Uses TRMM and multi-satellite precipitation, MODIS and AVHRR Land Cover, and Landsat topography	0.25° - 1° 3-hourly, daily, monthly 1979-2010	Rodell et al., 2004
MOD16	Normalised vegetation Index (NDVI) based model. Remote sensing observations from MODIS	1km globally 8-day, monthly 2000-2014	Cleugh et al., 2007 Mu et al., 2007 Mu et al., 2011 Fisher et al., 2011
METRIC – Mapping Evapotranspiration with Internalized Calibration	Based on an energy balance model, using Landsat data	30m globally, 2011-2016	Allen et al., 2007
GLEAM – Global Land Evaporation Amsterdam Model	Use the Priestly and Taylor equation, based on air temperature and net radiation	0.25°x0.25° Daily 1980-2016	Martens et al. 2017
ECMWF EartH2Observe	Water resource reanalysis data, created using an ensemble of 10 land surface and hydrological models	0.5°x0.5° Daily 1979-2012	Schellekens et al., 2017

GLEAM

The Global Land Evaporation Amsterdam Model (GLEAM v3, Martens et al., 2017) is a set of algorithms that estimate terrestrial evaporation and root-zone soil moisture from satellite data. GLEAM derives the different components of terrestrial evaporation, i.e. transpiration, bare soil evaporation, open-water evaporation, interception loss and sublimation (Miralles et al., 2011). Each grid cell, at 0.25°x0.25° resolution, is made of four different land-cover types – bare soil, low vegetation (e.g. grass), tall vegetation (e.g. trees), and open water (e.g. lakes and rivers). The fractions that are assigned to each of these land-covers are sourced from the Global Vegetation Continuous Fields product (MOD44B), based on observations from the Moderate Resolution Image Spectroradiometer (MODIS), and for open water the product of Tuanmu and Jetz (2014). Evaporative fluxes are calculated for each of the land-cover types separately and then aggregated to the scale of the grid boxes based on the fractional cover of each type. The Priestly and Taylor (1972) equation (see Equation 1) is used to calculate the cover dependent PET (mm/day) based on air temperature and net radiation, and this limits the number of spatially-varying surface fields that need to be specified and cannot be detected from satellite observations (Miralles et al., 2011) that other radiation-driven evaporation models require.

$$PET= \frac{1}{\lambda}\frac{s}{s+\ \gamma}\alpha(R_n\text{-}G)$$

where PET is potential evapotranspiration, λ is the latent heat of water vapour, α is a model coefficient (for daily calculation was determined to be 1.26 for freely evaporating surfaces; Priestly and Taylor, 1972; Stewart and Rouse, 1977), s is the slope of the saturation vapour density curve, γ is the psychrometric constant, R_n is net radiation, and G is soil heat flux.

ECMWF EARTH2OBSERVE

EartH2Observe “Global Earth Observations for Integrated Water Resource Assessment” is a project funded by the European Union (Schellekens et al., 2017). The overall aim is to contribute to the assessment of global water resources through the use of new earth observation datasets and techniques. Integrating earth observations, in-situ datasets and models, the project constructed a daily global water resource reanalysis dataset for the time period 1979-2012, with spatial resolution 0.5°x0.5°. A number of variables were outputted and form this dataset, and PET is among these. The dataset consists of an ensemble of 10 global hydrological and land surface models. These are: HTESSEL-CaMa, JULES, LISFLOOD, ORCHIDEE, PCR-GLOBWB, SURFEX-TRIP, SWBM, W3RA, WaterGAP3, and HBV-SIMREG. All models are forced with WFDEI (WATCH Forcing Data Methodology applied to ERA-Interim reanalysis; Weedon et al., 2014). A large ensemble of models was used to mitigate some of the uncertainties that are associated with using a single model from the simplification of the representation of spatially heterogeneous processes (Vrugt et al., 2005;

Gosling et al., 2010). However, this does not mean that some models do not perform better in specific locations, climatic conditions, or for specific variables than others.

TEMPERATURE BASED ESTIMATE OF PET

Finally, a temperature-based estimate of PET was constructed using a simple relationship between temperature and radiation, as devised by Oudin et al. (2005). The method was developed in order to combine simplicity and efficiency and make use of available atmospheric variables to represent evaporative fluxes at the basin level. Equation 2 gives this relationship.

$$\mathrm{PET} = \frac{R_e}{\lambda\rho}\frac{T_a+K_2}{K_1} \quad \text{if } T_a + K_2 > 0$$

$$\mathrm{PET} = 0 \qquad \text{otherwise}$$

Where PET is the rate of potential evapotranspiration (mm day^{-1}), R_e is extra-terrestrial radiation (MJ m^{-2} day^{-1}) depending only on latitude and Julian day, λ is the latent heat flux which is taken to be equal to 2.45 MJ kg^{-1}), ρ is the density of water (1000 kg m^{-3}), and T_a is the mean daily air temperature. Daily mean temperature was taken from the WFDEI forcing data (Weedon et al., 2014), and this is gridded with spatial resolution 0.5°x0.5°.

Oudin et al. (2005) evaluated the performance of a lumped rainfall-runoff model using values from 27 PET models and their impacts on streamflow simulation for 308 catchments in a wide range of climatic zones. The result of this study showed that very simple methods of calculating PET relying only on mean daily temperature and extra-terrestrial radiation (inferred from latitude and day of the year) are as efficient as more complex models such as the Penman-Monteith (Monteith, 1965).

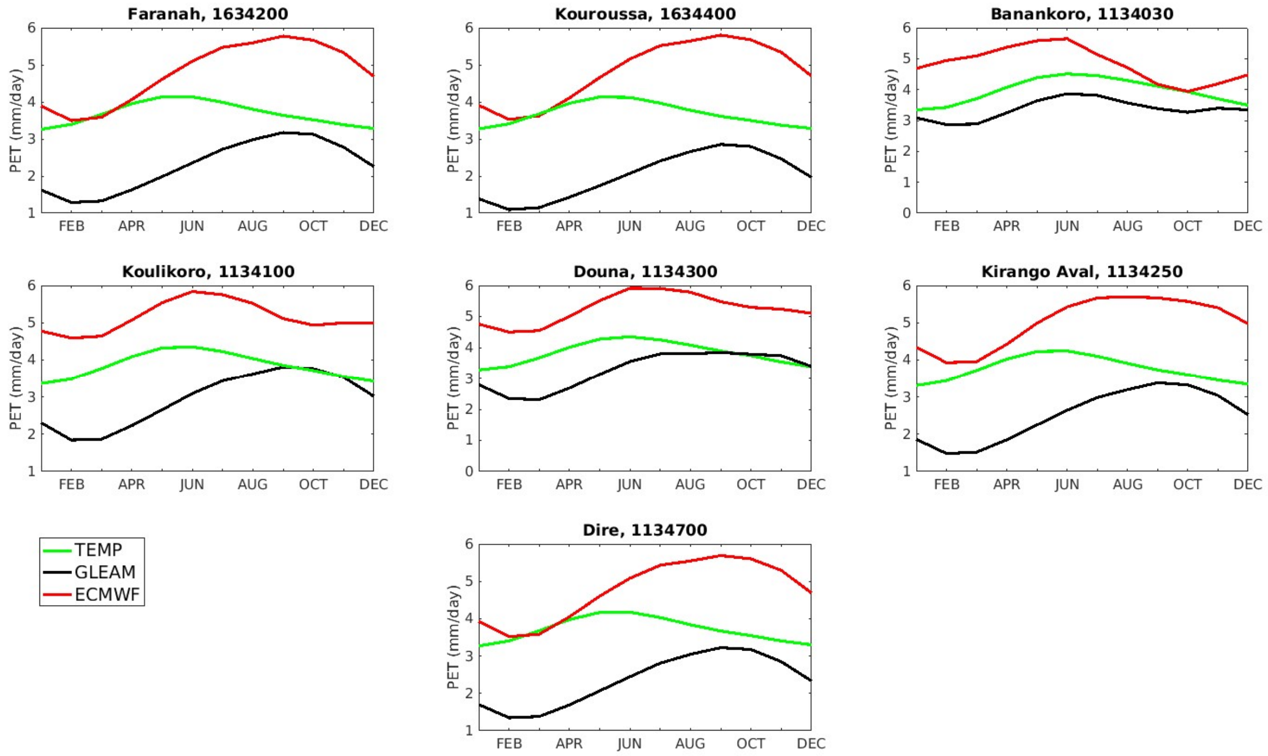


Figure 3.9. Mean monthly PET for each dataset for the time period 1980-2000, in each of the sub-basins in upstream to downstream order.

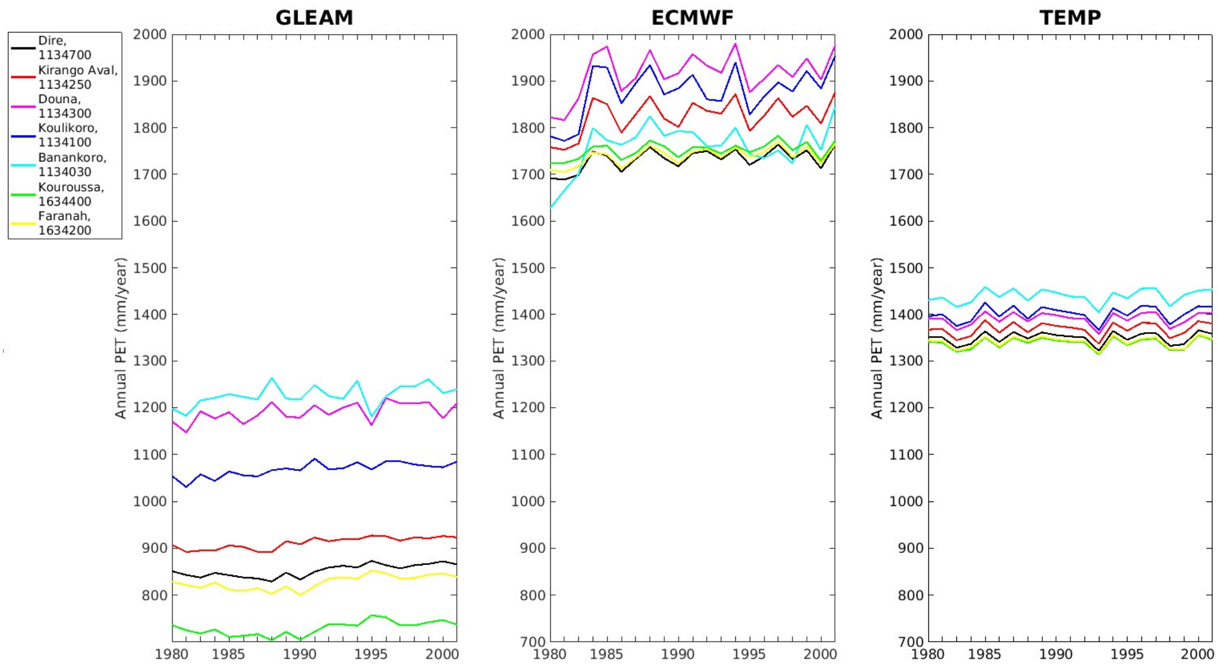


Figure 3.10. Interannual variability and differences between the seven sub-basins, for the three PET datasets used.

3.4. MODEL PERFORMANCE EVALUATION

The outputted river flow time series for each group of 10,000 model simulations from each of the sub-basins, as defined by their downstream gauging station, were evaluated against observed discharge at the seven stations available. In order to produce a robust evaluation of the model's performance across the Upper Niger basin with its spatially varying hydroclimatic characteristics, multiple performance evaluation criteria were used. This is so that the model's ability to capture a range of hydrologic responses can be assessed. Three model metrics were chosen. Nash Sutcliffe Efficiency (NSE, Nash and Suttcliffe, 1970) is used to assess the model simulations ability to capture the timing, shape and magnitude of high flows and peaks in the hydrograph. Percent bias (PBIAS, Gupta et al., 1999) is used to evaluate model performance in maintaining overall water balance in each of the sub-basins. Bias in low flow volumes (LFVBIAS, Yilmaz et al., 2008) was also calculated in order analyse the model simulations ability to capture low flows. These metrics will be discussed in the following sub-sections.

NASH SUTCLIFFE EFFICIENCY

NSE is used to assess the predictive power of hydrological models, and is defined as:

$$NSE = 1 - \frac{\sum_{t=1}^{n}(Qobs_t - Qsim_t)^2}{\sum_{t=1}^{n}(\overline{Qobs}_t - \overline{Qobs}_t)^2}$$

where Qsim is the simulated flow, Qobs is the observed discharge, $\overline{Qobs}$ is the mean of the observed discharge timeseries, and n is the total timesteps. NSE can range from $-\infty$ to 1, with NSE = 1 being a perfect match between the modelled discharge and the observations, and NSE = 0 meaning the model simulations are performing with the same skill as the mean observed flow.

There is much debate in the literature concerning the usefulness and strength of NSE as a performance metric (e.g. Krause et al., 2005; Schaefli and Gupta, 2007; Moriasi et al., 2007). The numerator of the equation above shows that smaller errors will become smaller, and larger errors will become larger. This is a major disadvantage of the criterion, as it leads to overestimation or underestimation of performance, depending on the model simulations that are being evaluated (McCuen et al., 2006; Jain and Sudheer, 2008; Gupta and Kling, 2011). Another of its disadvantages is that is uses the observed discharge mean in the calculation of variance. This is because in catchments that have highly variable discharges, it leads to overestimation of the performance of the model. However, while the limitations are recognised, NSE is a simple model metric which is easy to implement and has been used extensively in the literature that allows for comparison between sub-basins within this study and also with other similar hydrological modelling studies.

PERCENT BIAS

PBIAS measures the tendency of the simulated flow to be over or under predicted by the model compared with the observed data, and is calculated as:

$$PBIAS = \frac{\sum_{t=1}^{n} (Qobs_t - Qsim_t)}{\sum_{t=1}^{n} Qobs_t} X\ 100$$

where Qsim is the simulated flow, Oobs is the observed discharge, and n is the number of timesteps. The optimal value is 0. Positive values indicate the model is underestimating the bias in the flow, and negative values indicate an overestimation. This is a commonly used performance metric for quantifying water balance errors in model simulations (Moriasi et al., 2007) and therefore allows for intercomparison with other modelling studies that have used this as an evaluation criteria.

LOW FLOW VOLUME BIAS

The percent bias calculated for specific parts of the flow duration curve can also be calculated, as proposed by Yilmaz et al. (2008). To evaluate the model's capacity to represent low flows, the percent bias in low volume flows was also calculated, for flows that exceeded 70% on the flow duration curve.

$$LFVBIAS = \frac{\sum_{l=0.7}^{1} [\log(Qsim_l) - \log(Qsim_l)] - \sum_{l=0.7}^{1} [\log(Qobs_l) - \log(Qobs_l)]}{\sum_{l=0.7}^{1} [\log(Qobs_l) - \log(Qobs_l)]} X\text{-}100$$

A threshold value for low volume flows was calculated for each of the sub-basins which represents flows that have a flow exceedance of 70-100% on the flow duration curve. Each of the model simulations was compared with the observed discharge timeseries, and the percent of the time where the simulations over or under predict the observations below this threshold was calculated.
The use of flow signature indices is a way of separating the hydrograph and being able to target specific sections that are of interest. However, this metric is not timestep based and therefore does not look at the timing of the response of the sub-basins to the reduction in flows at the end of the wet season, which is an important aspect of water management. The flow duration curve is calculated to show the relationship between observed discharge magnitude and the exceedance probability of runoff in catchments (Ley et al., 2016; McMillan et al., 2017). It is a valuable hydrological signature of runoff variability and can be used for model performance evaluation and intercomparison with other hydrological modelling studies (Vogel and Fennessey, 1994; Yilmaz et al., 2008; Westerberg et al., 2011).

These three performance metrics were also combined in order to determine a behavioural ensemble of parameter sets. To do this, the scores calculated for each of the evaluation metrics were individually ranked in order of best to worst performance. These ranks were then summed for each of their corresponding model simulation to get a combined rank. Then these combined ranks were sorted in order from lowest to highest. Threshold values for each of the metrics were defined that model simulations needed to satisfy in order to be classified as behavioural. These thresholds are based on ranges that have been reported in the literature (e.g. Moriasi et al., 2007). These are NSE > 0.5, -10 < PBIAS < 10, and -10 < LFVBIAS < 10. Uncertainty bounds were also calculated within a Generalised Likelihood Uncertainty Estimation (GLUE) framework (Beven, 2006). This is an example of a methodology used to estimate the uncertainty of model predictions, and is able to include multiple metrics within its framework (Beven and Freer 2001b). The basic principle of GLUE is that we lack knowledge of how environmental systems work exactly, and therefore this cannot be represented perfectly with our hydrological model structures. Due to this inability, there will always be several different models that can perform with a similar level of performance, and these behavioural models are called equifinal. The methodology uses results that are expressed as a probability distribution, and analyses how accurate these representation are of the uncertainty.

The ability of the behavioural simulations to capture the inter- and intra-annual variability was also evaluated. This can be characterised by different properties (Gudmundsson et al., 2012). The mean annual value is a measure of the water balance in a catchment and can give insight into the long-term water balance. The difference in the highest to the lowest simulated flow value, or the amplitude, is a measure of how pronounced the seasonal cycle of model simulations is when compared to the observed flow.

To measure the ability of the 'behavioural' simulations to capture the mean annual runoff in the sub-basins, the relative bias was calculated:

$$\Delta\mu = \frac{\Delta\mu_{sim} - \Delta\mu_{obs}}{\Delta\mu_{obs}}$$

Where $\Delta\mu_{sim}$ and $\Delta\mu_{obs}$ is the mean annual cycle of behavioural simulations and observations, respectively. The optimal value is $\Delta\mu = 0$, and absolute values of 1 is a deviation of 100%. To measure the ability of the 'behavioral' simulations to capture the amplitude of the observed annual cycle the difference in standard deviation was calculated:

$$\Delta\sigma = \frac{\Delta\sigma_{sim} - \Delta\sigma_{obs}}{\Delta\sigma_{obs}}$$

Where $\Delta\sigma_{sim}$ and $\Delta\sigma_{obs}$ is the standard deviation of the simulations and the observations mean annual cycle, respectively. The ability of the model simulations to capture the monthly means when compared with that of the observations was also calculated.

4. RESULTS

DECIPHeR was set up for the Upper Niger basin with a total catchment area of 371,920km^2. Sub-basin area ranged from 3,214.14km^2 to 238,350.7km^2. After the digital terrain analysis, there were 1843 HRUs for the whole catchment area, with 40-558 in each sub-basin. Sub-basins will be referenced by their downstream gauging station codes throughout this section, and this information can be found in Table 3.7. HRU areas varied from 0.009025km^2 to 1810.803km^2, with a median of 82km^2 and mean of 183.782km^2. One simulation for the 21-year time period for the whole Upper Niger Basin takes approximately 3 minutes to run, outputting a simulated flow time series for the seven basins used in this study.

4.1. INITIAL RESULTS

To provide a benchmark of model performance, DECIPHeR was run for the Upper Niger basin using the model structure and parameter ranges used in Coxon et al (2018) (described in Section 3.2) and using a single forcing dataset consisting of MSWEP global rainfall data (Beck et al., 2018) and GLEAM global PET data (Martens et al., 2017). The main objective of this project was to test model capacity to be used in large and data sparse river basins, and the initial results can be used to guide future model developments and evaluate where the model may require improvements.

The three performance metrics were calculated over the 21 year time period for each of the 10,000 simulations for each sub-basin. This number of parameter sets was chosen as it provides a representative sample of the parameter space and a reasonable model run time (approximately 6 days for 10,000 parameter sets). Figure 4.11 shows the uncertainty bounds for the top performing 100 simulations in all sub-basins compared with the observed discharge at each of their downstream gauging stations. The simulated discharge is doing reasonably well at capturing the timing and the shape of the peaks, increasing in response to the high rainfall of the wet season, however it is hugely overestimating the volume of these flows, approximately three times the observed in all sub-basins. This is indicated with negative NSE and very large PBIAS scores. For example, the best NSE scores in the sub-basins range from -1.62 at 1134300, to -50.62 at 1134700. The best PBIAS scores range from 203.19 at 1634200, to 456.35 at 1134250. The model is performing well in terms of capturing low flow volumes, with some simulations gaining LFVBIAS scores close to zero, however, these are not the same simulations that are achieving the upper NSE and PBIAS scores, as these simulations are overestimating the magnitude in sub-basins and missing the timing of the recession in flow volumes at the end of the wet season.

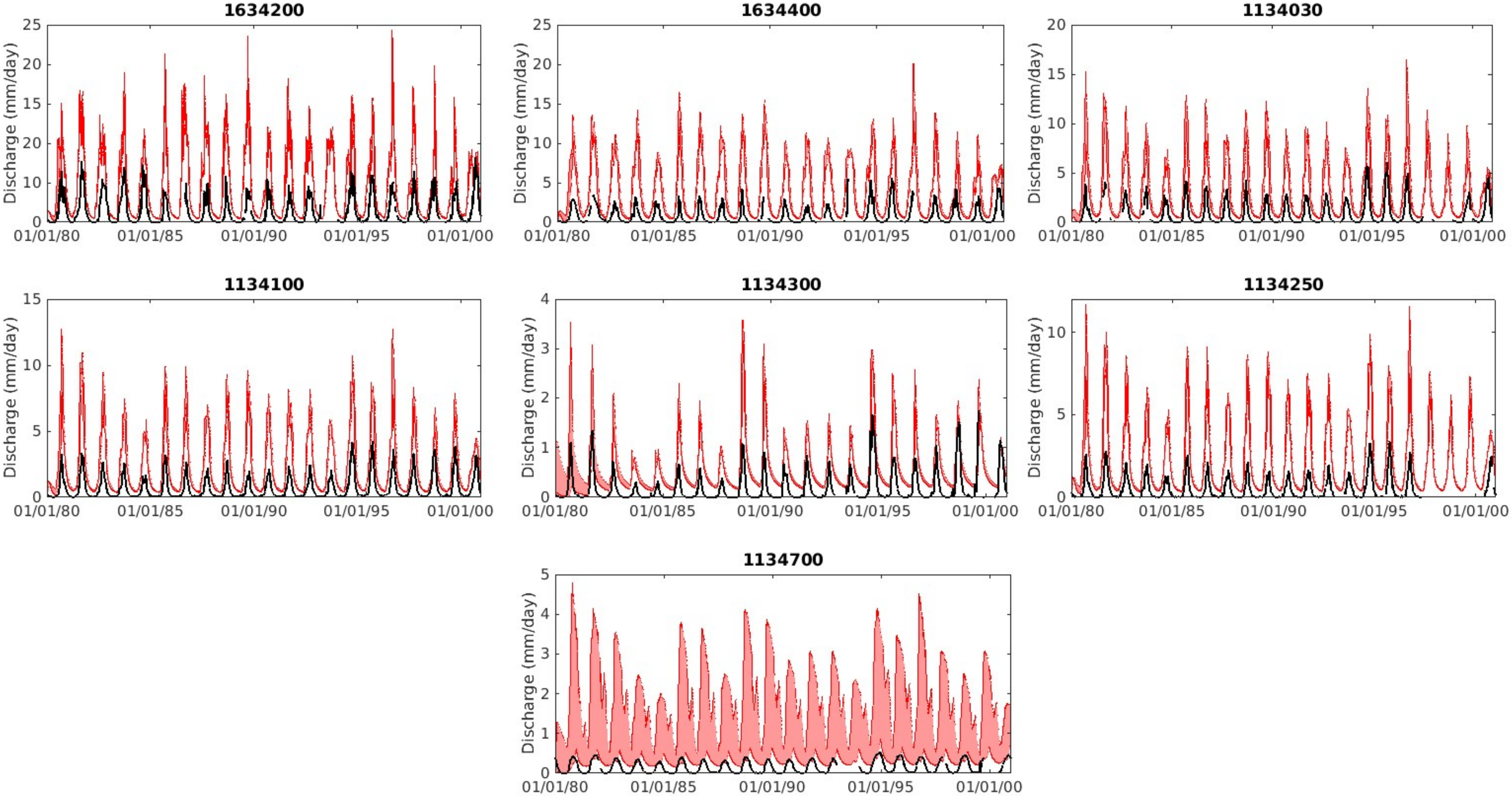


Figure 4.11. Uncertainty bounds for top 100 performing simulations in all sub-basins using the initial model structure and parameterization with GLEAM PET forcing data.

This overestimation of peak flow magnitudes is seen for all model simulations for all seven gauging stations in the model domain.

To investigate this further, the water balance for each of the sub-basins was calculated using the input data used for the initial model simulations. This is summarised in Table 4.11. The gauging stations are given in upstream to downstream order.

There is a large surplus of water in the upstream catchments, which indicates that with the current input data, the water balance cannot be closed. DECIPHeR maintains mass balance, meaning that the volume of water entering the catchment from rainfall can only be lost via evaporation or leaving the catchment via the downstream outlet, and therefore water cannot be taken from the system at the rate required in order for simulated flow to accurately predict the magnitude of the peaks. Water loss could be due to a number of factors such as uncertainties in the input data, inter-catchment groundwater flows and/or human impacts.

Table 4.11. The water balance for each sub-basin, as defined by their downstream gauging station, is calculated from the mean annual precipitation, taken from the MSWEP global data, mean annual PET, taken from the GLEAM global data, and mean annual discharge measured at the gauging stations. Positive values indicate where there is a surplus of water, and negative values indicate where there is a deficit.

GAUGE	MEAN ANNUAL PRECIPITATION (m)	GLEAM MEAN ANNUAL PET (mm)	MEAN ANNUAL Q (mm)	WATER BALANCE (P-PET-Q, mm)
1634200	2295.5	684.3	451.2	1160
1634400	2102.9	805.2	248.8	1048.9
1134030	1555.9	882.1	266.4	407.4
1134100	1094.1	928.0	246.4	-80.3
1134300	710.0	1064.5	63.0	-417.5
1134250	700.3	1194.4	149.8	-643.9
1134700	319.9	1227.0	50.9	-958

For the downstream sub-basins, the water balance is negative. This suggests that there is enough PET currently estimated in the input data, however it is not being used in the model to correctly simulate discharge as there is still substantial over-predictions of flow volumes in these catchments. This suggests that there may be issues with the current model structure and/or the parameterisation. To explore this further, the proportion of PET that is being taken as actual evapotranspiration (AET) for the whole of the simulation time period was investigated for all 10,000 model simulations at each of the gauges.

Figure 4.12 shows the relationship between the proportion of PET that is utilised and the parameter SRmax, which controls the amount of evapotranspiration that can occur within the basin, as it defines how deep the soil root zone is, and therefore how much water can be held in this store. For the upstream gauging stations, i.e. 1634200, 1634400 and 1134030, more than 70% of the input PET is being taken as AET for most of the simulations, and is as high as 90% for the highest SRmax values. This is showing that the model is losing water from the system at almost full potential rate. For the downstream gauges, where the water balance is negative, i.e. 1134100, 1134300, 1134250 and 1134700, a much smaller proportion of PET is being used as AET by the model, approximately ranging between 20 and 50%. This is interesting, as noted above, the negative water balance suggests that there is enough PET, however it is not being used.

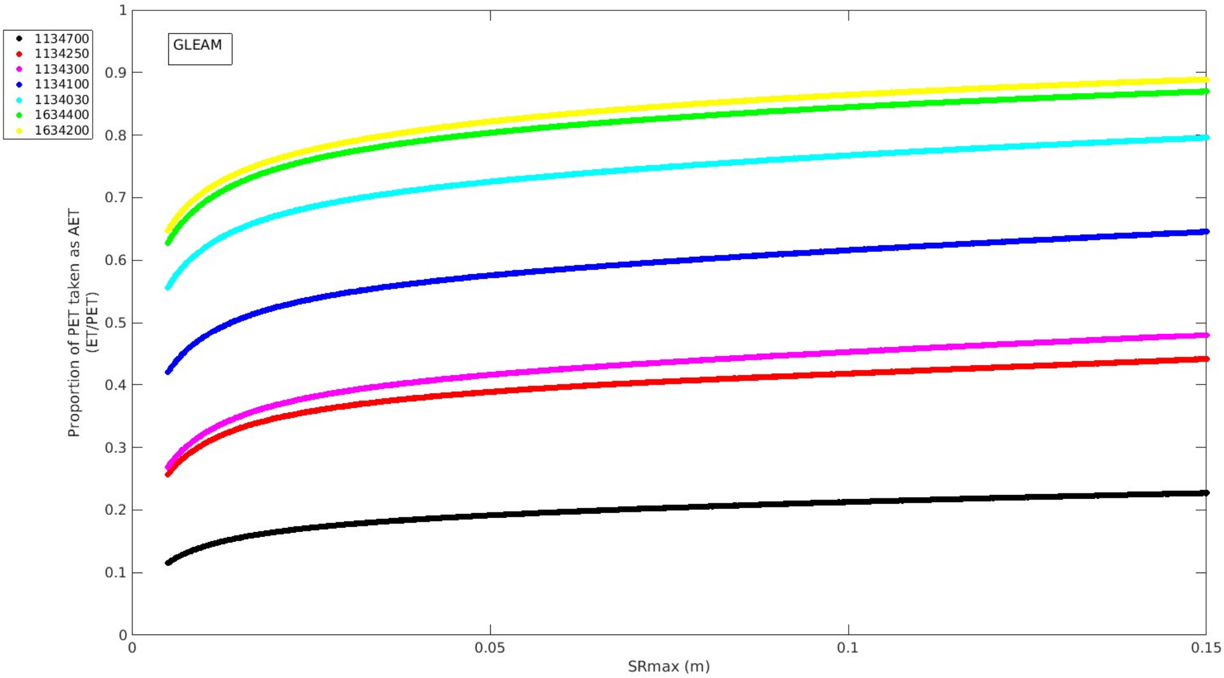


Figure 4.12. The proportion of input PET that is taken as actual evapotranspiration (AET) from each sub-basin, as defined by their downstream gauging station, for each of the 10,000 model simulations for the whole of the simulation period. This is controlled by SRmax, the parameter that defines how deep the soil root zone is, and therefore how much water can be stored here.

To further investigate the way that input PET is being treated within the model, the relationship between AET and PET, and storage in the soil root zone is shown in Figure 4.13 for one sub-basin, 1134100, for one simulation using one set of parameters. This sub-basin was chosen as it is one of the downstream sub-basins where the calculated water balance is negative, suggesting there is enough PET. However, it is not being used as AET at a high enough rate to accurately predict river flows. This simulation was chosen as it was the best performing when evaluated using three performance metrics. Here, it can be seen that AET can only be taken at almost full potential rate during the wet season when the soil root zone has been replenished. This figure may also provide insight into why the model simulations seem to perform better during the low flows (seen in Figure 4.13D). There seems to be almost no rainfall between November and March, meaning there is no input into the soil root zone. Consequently, the soil root zone dries out over this period through evaporation directly from this store and drainage to the saturated zone, and therefore provides no direct overland run-off into streamflow Figure 4.3 also indicates that there is potential for the soil root zone to store more water if it were larger, which in turn would allow for more evaporation to take place at the full potential rate.

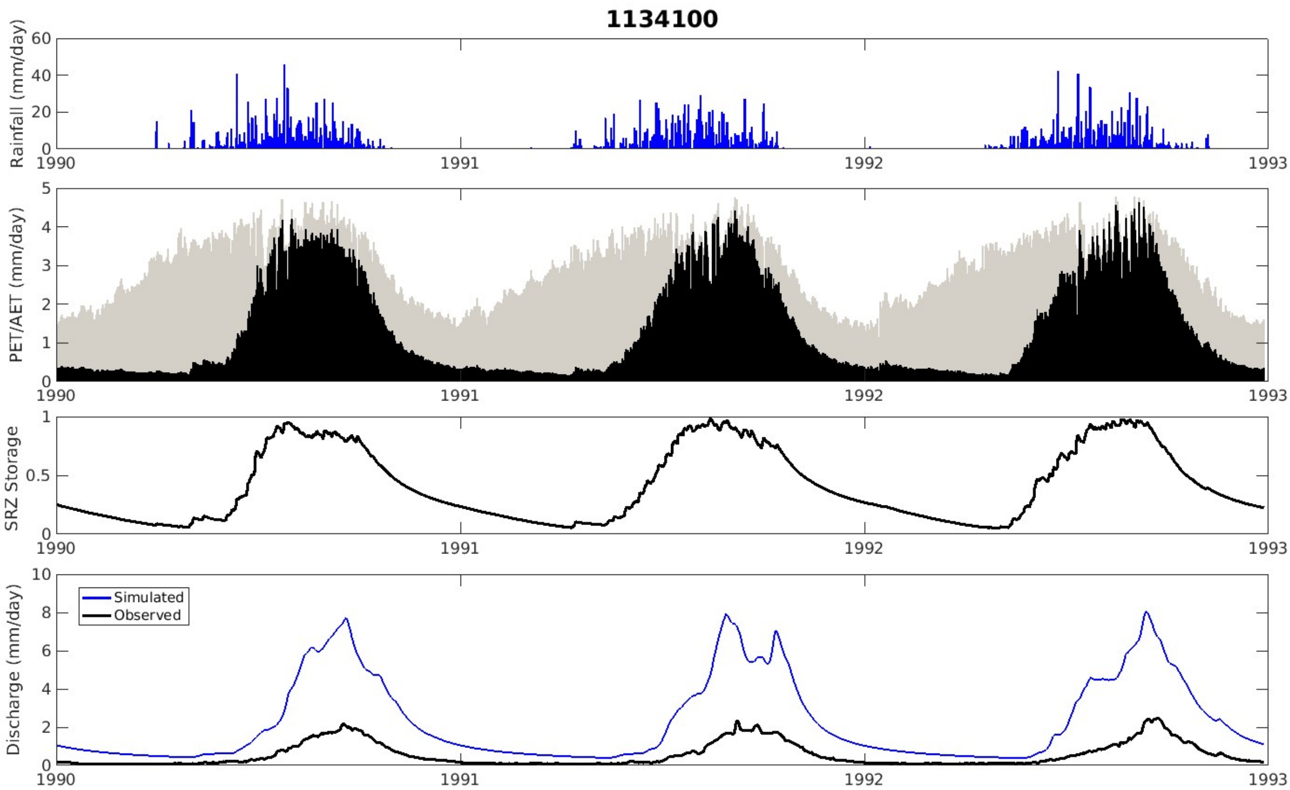


Figure 4.13. For a chosen parameter set for one year of the simulation in one sub-basin: A. Mean daily precipitation, B. PET and AET removed from the basin by the model, C. Proportion of the soil root zone that is being used for storage, D. Comparison of the simulated flow with the observed discharge at gauging station 1134100.

4.2. INPUT DATA AND PARAMETERIC UNCERTAINTY ANALYSIS

The initial results indicate that there are some fundamental processes in the Upper Niger basin that are not currently being represented with the first set of model simulations. The inconsistencies between the water balance data for each of the sub-basins and simulated river flows indicated that this is potentially due to errors in the input data, parameter sets, model structure, but it is likely a combination of these. To explore the uncertainty due to input data, two more PET datasets were used to force the model simulations – one global PET product (ECMWF) and a dataset constructed based on a simple relationship between mean daily temperature and latitude (referred to as TEMP). These datasets are described in Section 3. Table 4.12 summarises the mean annual total PET for the three datasets and the water balance for each of the sub-basins. From this comparison, it is clear that there is large uncertainty in the input data used. ECMWF PET totals are considerably higher than GLEAM, approximately 900 - 1000 mm/year higher in most of the sub-basins. The variability within the datasets themselves is very different also. For GLEAM, there is approximately 550 mm/year difference between the headwater sub-basin, and the most downstream gauge. For ECMWF, there is only approximately 200 mm/year difference between the sub-basins with the highest and lowest mean annual totals. This difference is even lower for the PET estimate based on daily mean temperature, at 100 mm/year. However, this summary table only shows the mean annual totals across the whole of the sub-basins.

Table 4.12. Comparison of the mean annual PET estimates and water balance for the sub-basins defined by their downstream gauging station for three PET datasets. GLEAM and ECMWF are global products, and TEMP refers to PET data that was calculated based on a simple relationship between mean daily temperature and latitude.

GAUGES	PRECIP (mm/year)	Q (mm/year)	ECMWF PET (mm/year)	WATER BALANCE	TEMP PET (mm/year)	WATER BALANCE
1634200	2295.5	451.2	1758.9	85.4	1339.3	505
1634400	2102.9	248.8	1734.0	120.1	1342.2	511.9
1134030	1555.9	266.4	1735.7	-446.2	1351.5	-62
1134100	1094.1	246.4	1827.8	-980.1	1370.7	-523
1134300	710.0	63.0	1888.6	-1241.6	1404.0	-757
1134250	700.3	149.8	1922.8	-1372.3	1391.8	-841.3
1134700	319.9	50.9	1769.1	-1500.1	1438.9	-1169.9

Further analysis of the difference between the PET datasets was conducted by looking at the water balance in each of the sub-basins. Figure 4.14 shows the mean monthly water balance in three sub-basins. Where the rainfall is higher than PET, there is a surplus of water within the catchment, and where the PET is higher than the rainfall there is a deficit. All PET datasets produce a deficit during the dry season months, and this may indicate why model simulations perform well during low flows. However, during the wet season, GLEAM and TEMP estimate PET to be lower than ECMWF and produce a large surplus in the water balance in sub-basins. ECMWF estimates mean monthly PET to be higher, and in sub-basin 1134250 produces a negative water balance throughout the year. In sub-basin 1634200 (the headwater catchment of the basin) monthly mean PET for ECMWF is higher than both other datasets but during the wet season mean monthly precipitation is far higher. This is to be expected as this particular sub-basin is in the tropical Guinea Highlands, however, the surplus of water seems to be so large that with the current model parameterisation and structure, simulations with 'good' model performance cannot be obtained.

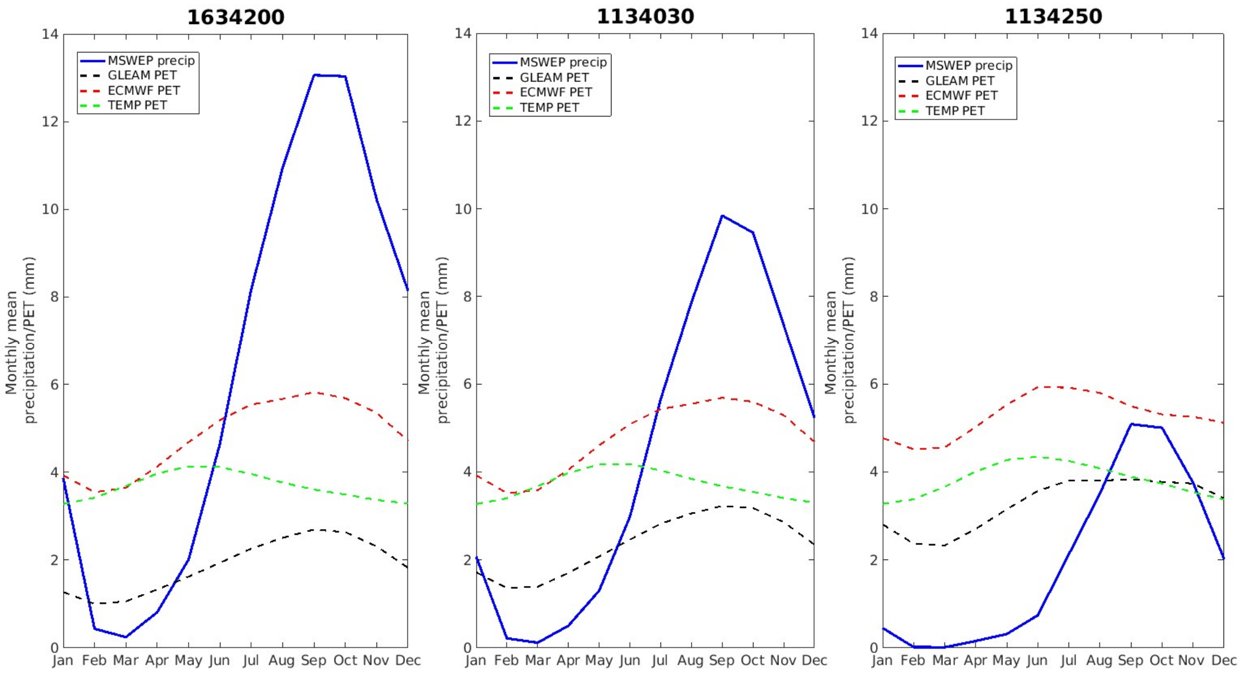


Figure 4.14. Comparison of the mean monthly water surplus/deficit in 3 sub-basins, as defined by their downstream gauging station, when being calculated using different PET estimates.

In addition to the mean annual totals and mean monthly water balance, the spatial pattern and variability of the three PET datasets was also analysed. Figure 4.15 shows this spatial variation of the mean daily PET across the Upper Niger basin. These spatial maps really highlight how different these different datasets are, and really emphasises how large input data uncertainty can be. This is particularly important in data scarce locations, such as the Upper Niger basin, where global input datasets cannot be extensively evaluated against observations or field measurements. In Figure 4.15, the spatial pattern for GLEAM is the inverse of that found in the MSWEP precipitation data, i.e. estimates are lower in the South-West headwaters and incline as the river moves downstream and northward. The variability for TEMP is similar in terms of starting lowest in the south-west and gradually increasing towards the downstream outlet. This is to be expected, as this dataset, as described above, has been constructed from a relationship between daily mean temperature and latitude (Oudin et al., 2005b), and temperatures are higher in the downstream region of the Upper Niger basin, where the river flows through the Sahelian belt. The variability in space for ECMWF PET is very different in comparison to GLEAM and TEMP. There are some similarities with lower estimates in the south and increasing downstream, however the lowest estimates are found at the basin outlet, and this is inconsistent with the other datasets.

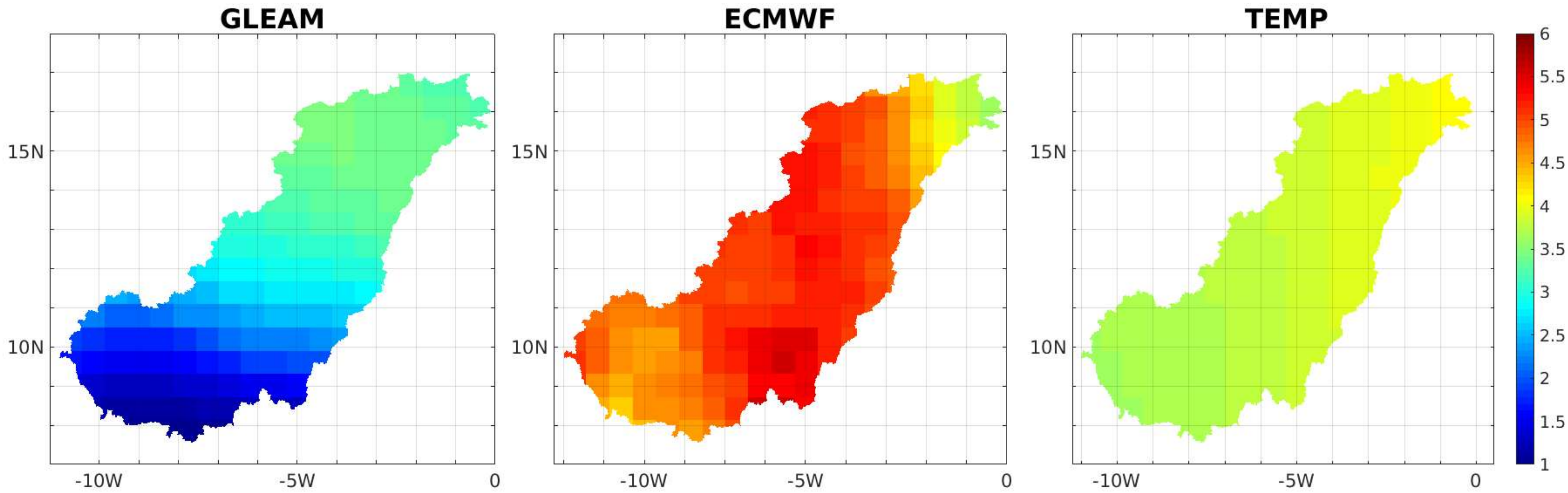


Figure 4.15. Spatial variability (mean mm/day) for three gridded PET datasets, with spatial resolution 0.5°x0.5°, for the Upper Niger basin.

DECIPHeR was run again with the same model structure and 10,000 parameter sets as the initial model set up, but with ECMWF and TEMP as PET input data. Model performance was evaluated using the same three metrics that were used to measure the performance of the first 10,000 simulations. Comparison of model performance for the whole ensemble of simulations using these metrics is summarised in Figure 4.16. For NSE, model simulations using ECMWF as PET input, performance is considerably improved, in terms of both absolute scores and overall interquartile range across the simulations. This is likely due to NSEs bias towards peak flows. As ECMWF has the highest estimates of PET, this will reduce simulated flow more than the other two PET input data sets. This leads to modelled discharge during high flows to be lower than in the initial model simulations, and therefore closer to gauged river flow, and obtaining better NSE scores.

A similar result is gained when measuring model performance with PBIAS. As above, simulations using ECMWF PET input data are improved in comparison with those using GLEAM and TEMP. This is again likely due to the way in which PBIAS is calculated. It measures the likelihood that model simulations are over/under predicting observed flows. As ECMWF estimates PET at a much higher rate per day, simulated flows are reduced in comparison to those produced by the initial simulations using GLEAM, and therefore closer to observations. TEMP PET also improves model performance slightly, mainly by reducing the interquartile range of performance scores.

For the LFVBIAS, model performance is very similar for simulations using each of the PET datasets. This is likely due to the climatic conditions of the Upper Niger basin and the model dynamics of DECIPHeR. For much of the basin, outside of the wet season there is very little rainfall, if any at all. Therefore, there is no input into the model domain for these areas during this period. This will create a drying of the model domain during this dry period, reducing simulated flows at each of the sub-basin gauging stations, and therefore performance when evaluated against observations obtains 'good' scores.

Figure 4.16 also shows the difference in performance between the sub-basins. When measuring performance with NSE and PBIAS, ECMWF PET data has improved scores in 1134030 and 1134100 the most. The range of scores for these sub-basins is much smaller in comparison with the other sub-basins. Model performance has been drastically improved in sub-basin 1134700, however it is still obtaining very low metric scores, and the range in values is very large.

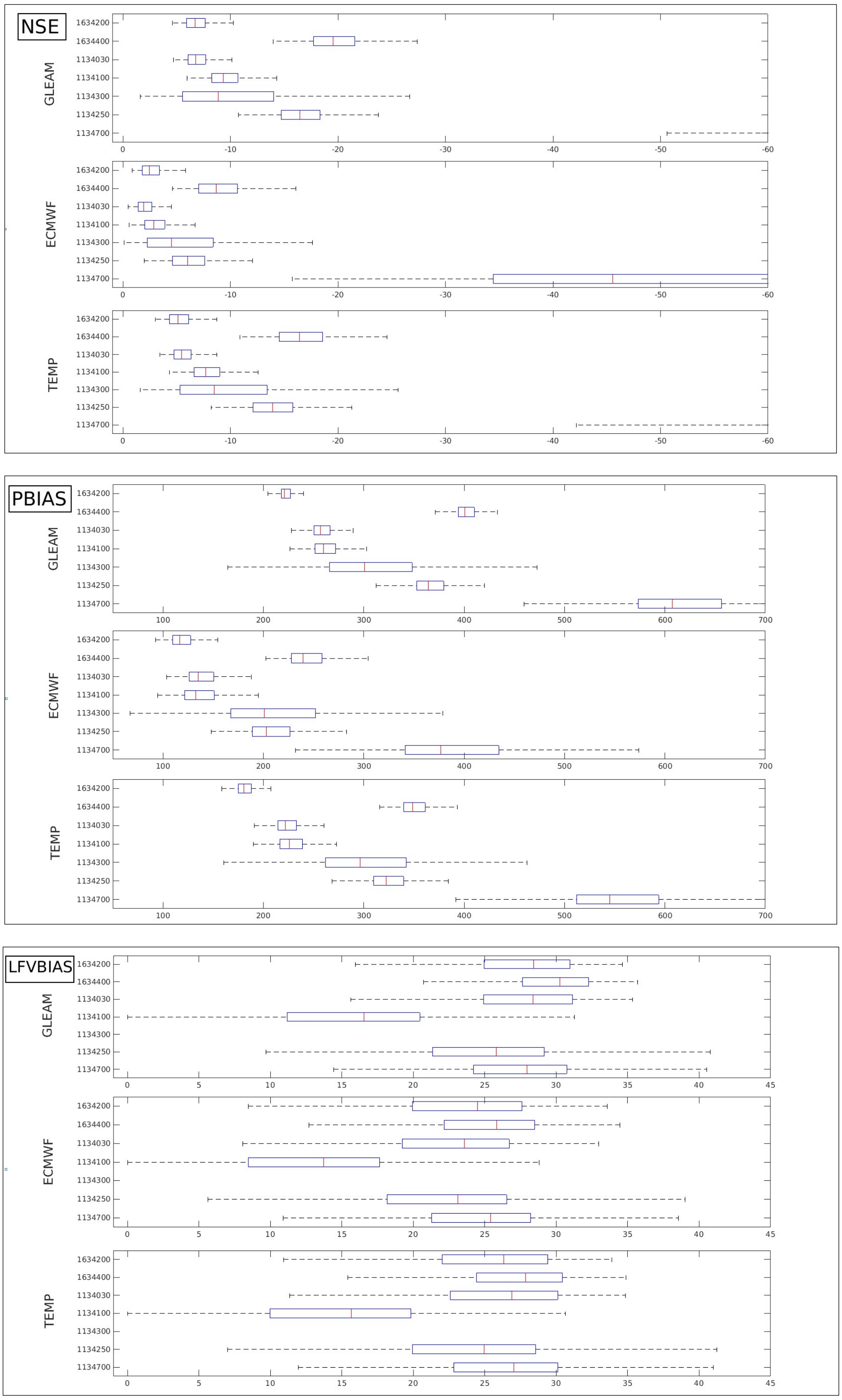
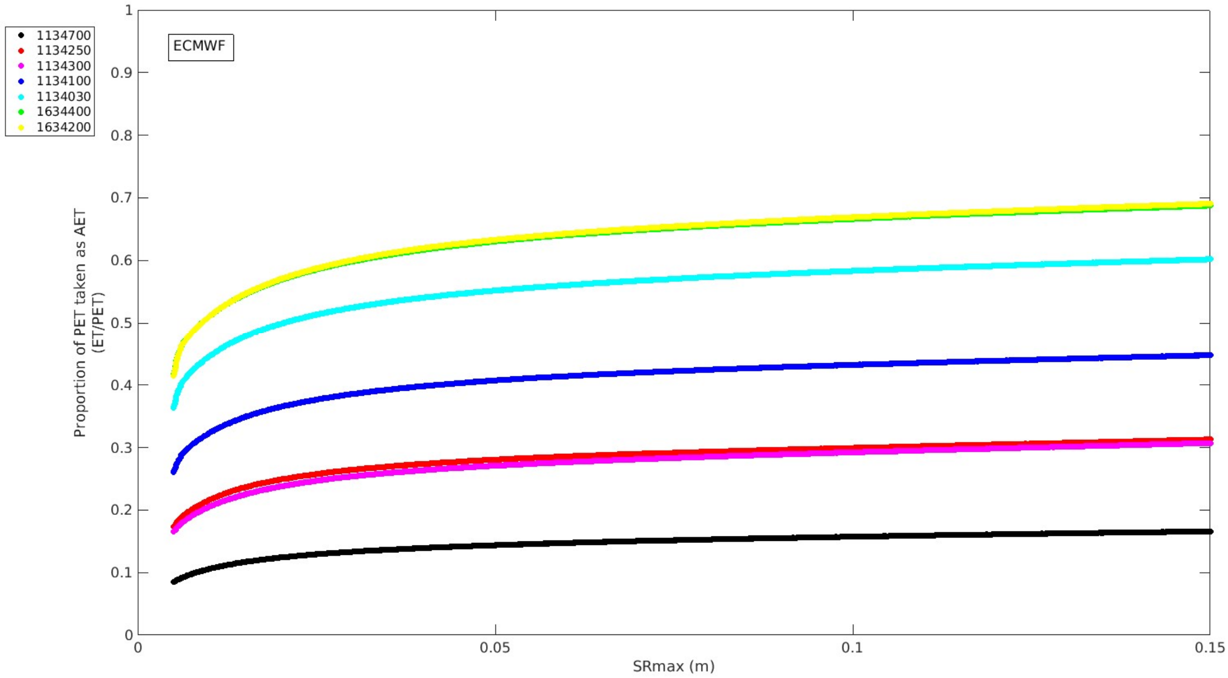


Figure 4.16. Comparison of model performance for all 10,000 model simulation, using the initial model set up with three different PET input datasets, for three performance metrics, Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and low flow volume bias (LFVBIAS).

An evaluation of how much AET is being taken in each of the sub-basins using the ECMWF and TEMP PET data was also need as a comparison with simulations using GLEAM. Figure 4.17 shows the proportion of PET that is utilised as AET for each model simulation for the whole 21-year period in each sub-basin, using ECMWF and TEMP data. When compared with Figure 4.12, it can be seen that the use of different PET datasets has a large effect on the proportion of input PET that is realised as AET in the basin. The largest difference with the initial model simulation can be seen in the upstream sub-basins of 1634200 and 1634400. The proportion of PET taken as AET when using GLEAM in these sub-basins ranged between 70-90% depending on the value of SRmax. However, this range is reduced to 50-70% when using ECMWF and TEMP. This reduction in the proportion of PET used as AET in model simulations indicates there is potential for reducing simulated high flow magnitudes, and therefore improving overall model performance.



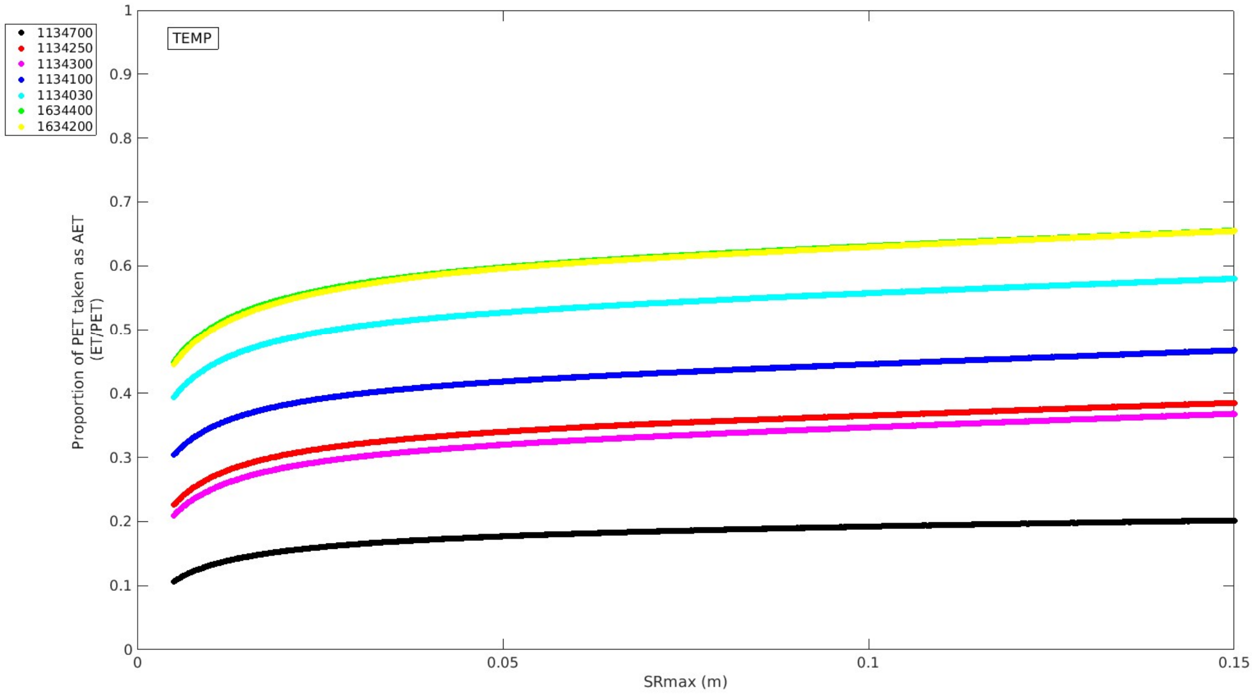


Figure 4.17. Proportion of PET that is used by DECIPHeR as AET for the 21-year simulation period, all 10,000 model simulations, at each sub-basin as defined by their downstream gauging station. This is compared with SRmax, as this is the parameter that controls the amount of evapotranspiration that can occur at each model simulation timestep.

Figure 4.12 and 4.17 show that the rate of evapotranspiration is completely dependent on the parameter SRmax, which controls the size of the soil root zone (in metres). In the initial model simulations, the upper and lower boundary for SRmax from which parameter sets were sampled was set at 0.15. This was increased to 0.5 and DECIPHeR was run with the same initial model structure, for 10,000 simulations, using each of the PET datasets. Model performance for each set of model simulation was evaluated using the same three performance metrics as with the initial model structure and parameters. Figure 4.18 shows the comparison in performance for one representative simulation at sub-basin 1134100 using the initial model and with SRmax increased, using each of the PET input datasets. Model performance is improved when using each of the PET products but is most notable for ECMWF. Looking at the hydrograph it is clear that the magnitude of the peak flows has been reduced, and NSE scores increased from -0.58 to 0.37 confirms this. The timing and shape of the peak is also improved for the ECMWF simulations. However, the timing of the recession of the hydrograph and volume of low flows after the wet season is not predicted correctly. This is also seen when using GLEAM and TEMP PET too.

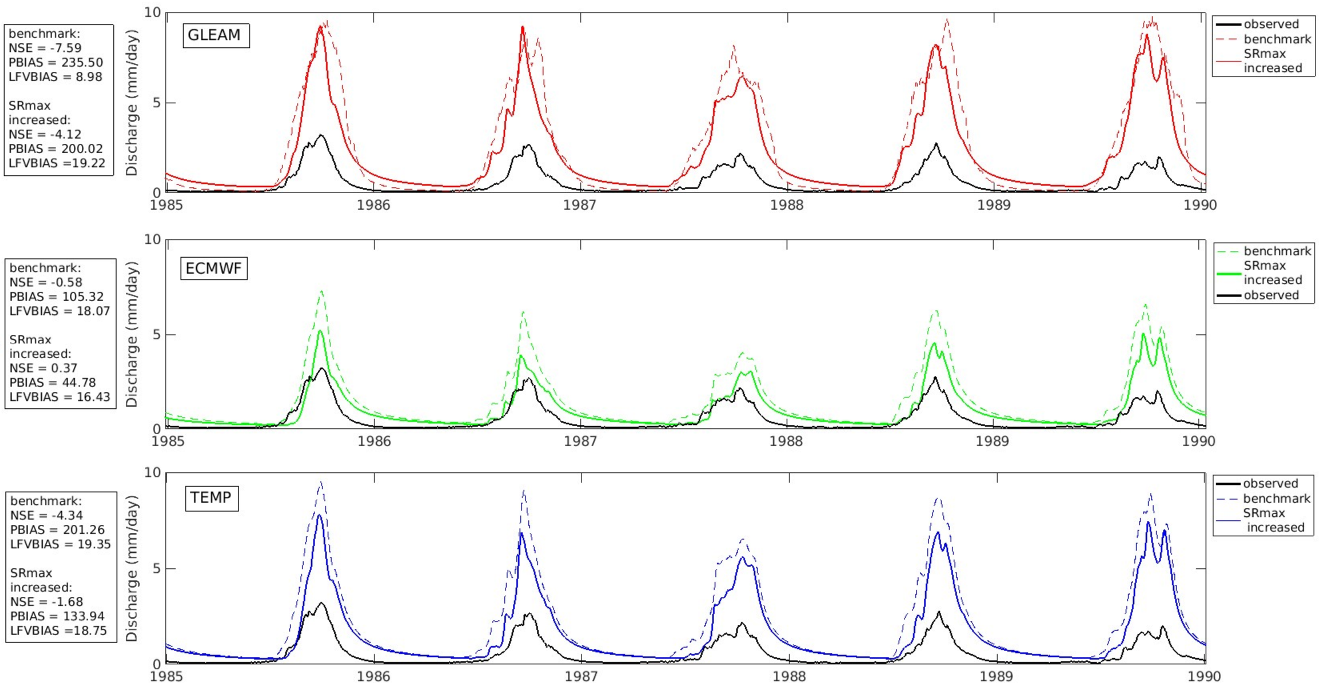


Figure 4.18. Comparison of initial model performance with model simulations where SRmax has been increased for 1985-1990, 5 years of the 21-year simulation period. One simulation at gauging station 1134100 is used for each of the model setups. Individual performance metric scores are given for each of the PET input datasets. 'Benchmark' refers to the initial model formulation.

The effect of this increase in SRmax on AET and storage in the SRZ is assessed in Figure 4.19 for one year of the 21-year simulation period, for one simulation for each of the PET datasets, in sub-basin 1134100. The pattern of AET that is used in the model simulation is similar in GLEAM and TEMP, as there is an almost consistent amount of AET during the dry season and then a rapid increase during the wet season to almost full potential rate. Looking at the AET in comparison with PET for ECMWF, it seems that there is much potential for more evaporation to be taken at a higher rate. In comparison with GLEAM and TEMP, during the wet season AET is being taken at a much lower rate, approximately half of the potential rate. This is likely due to the way that evaporation is taken from each of the HRUs in the model domain. In the current model structure AET is taken as a proportion of PET which is completely controlled by the storage in the soil root zone at each time step. For example, if the soil root zone is 50% full, AET is 50% of the PET input at that timestep. So, although SRmax controls the amount of water that can be held within the soil root zone as storage, if there is not enough rainfall within a sub-basin to fill this soil root zone store, evaporation cannot be taken at its potential rate.

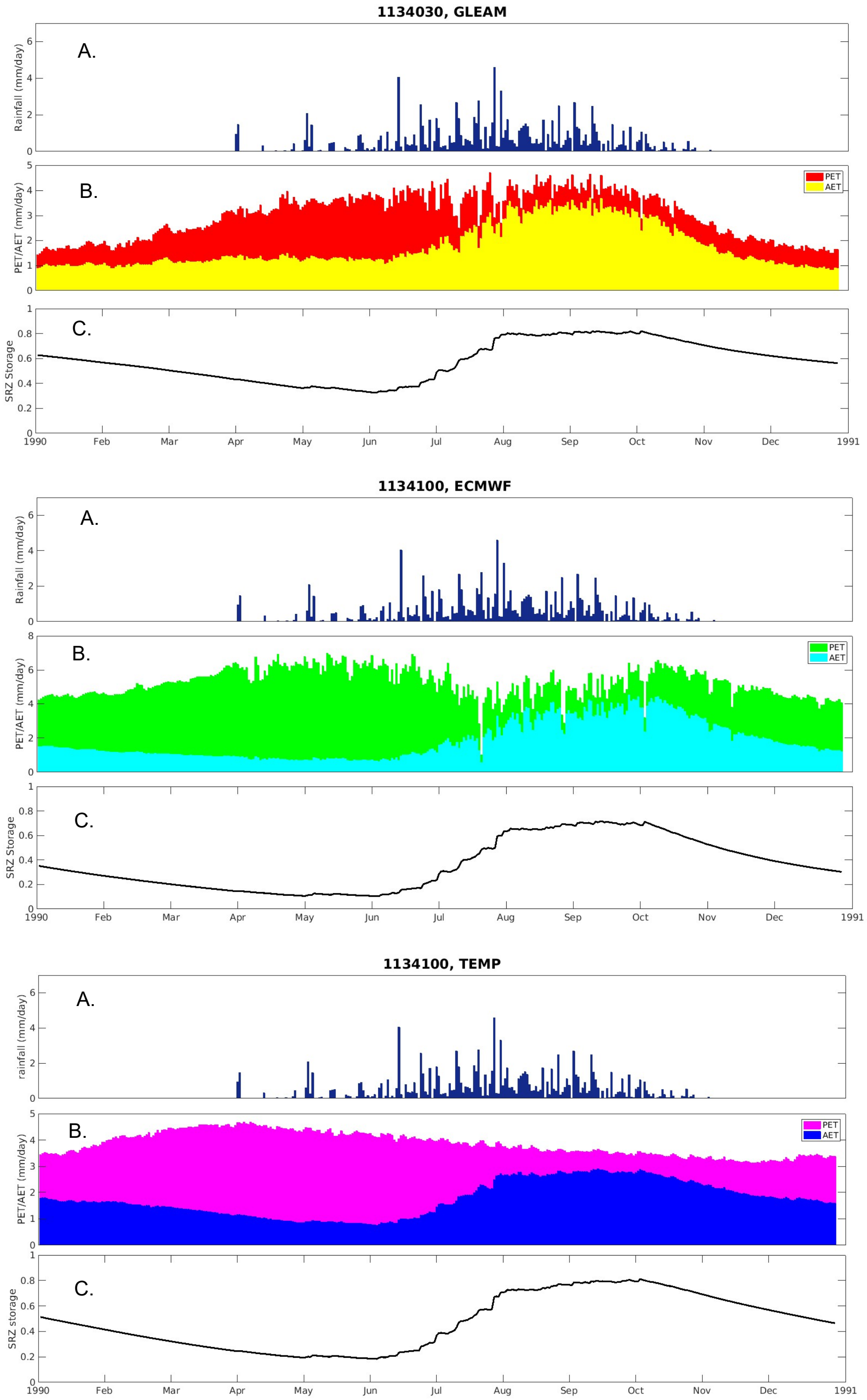


Figure 4.19. For a representative parameter set for one year of the simulation in one sub-basin: A. Mean daily precipitation, B. comparison of input PET (for each PET product) and AET removed from the basin by the model, C. Proportion of the soil root zone that is being used for storage.

4.3. MODEL STRUCTURAL UNCERTAINTY ANALYSIS

The results presented in the previous sub-sections provide evidence that there is a misrepresentation of the key processes that are occurring in the Upper Niger basin. The inconsistency between the water balance in the sub-basins, and the simulated discharges points to a misrepresentation of evaporation in the basin. Model performance was improved by changing the parameterisation to include a larger soil root zone across the basin, but this did not allow for 'good' performance to be obtained with the evaluation metrics used. Therefore, changes to the model structure are needed to improve model simulations in this catchment. A modified model structure was implemented that includes evaporation being taken from the saturated zone (see Figure 3.5). How this routine is calculated is described in more detail in Section 3.

4.3.1. OVERALL MODEL PERFORMANCE

The modified model structure was run with the same 10,000 parameter sets used in the initial model simulations and with the same 10,000 simulations with SRmax increased, using each of the PET input datasets. Model performance was calculated at each of the gauging stations of the sub-basins using each of the evaluation metrics. To look more closely at how the saturated zone evaporation routine has affected model performance in each of the sub-basins, thresholds of 'acceptable' model performance for each of the metrics was defined. These are NSE>0 (i.e. simulations performing better than the mean climatology of the catchment), -20 < PBIAS < 20 (i.e. model simulations are under/overpredicting observed flows for less than 20% of the time period), and -20 < LFVBIAS < 20 (i.e. low flow volumes greater than 70% on the FDC are being under/overpredicted for less than 20% of the simulation period). These threshold values are taken from Moriasi et al. (2007).

Figure 4.20 summarises the percentage of simulations in each of the sub-basins, when using each of the input PET datasets and each of the model formulations, that have a score greater than the threshold value for each of the three model performance metrics. Simulations when evaluated with LFVBIAS manage to obtain acceptable scores in all sub-basins, with all PET, and all model formulations. This shows that DECIPHeR is good at simulating low flows across the Upper Niger basin. For NSE, model performance in sub-basin 1134300 is improved with satevap added to the model structure when using all three PET input. However, it is only model simulations using ECMWF and with the satevap routine calculations where other sub-basins are also gaining acceptable model performance scores (i.e. 1134030, 1134100, 1134250). A similar result is found for when PBIAS is used to evaluate simulation performance. Table 4.13 also summarises the performance scores calculated for each model formulation, in each sub-basin with all PET inputs. In this table, only the highest scoring simulation is given, and where highlighted, the simulation is over the defined acceptable performance threshold values.

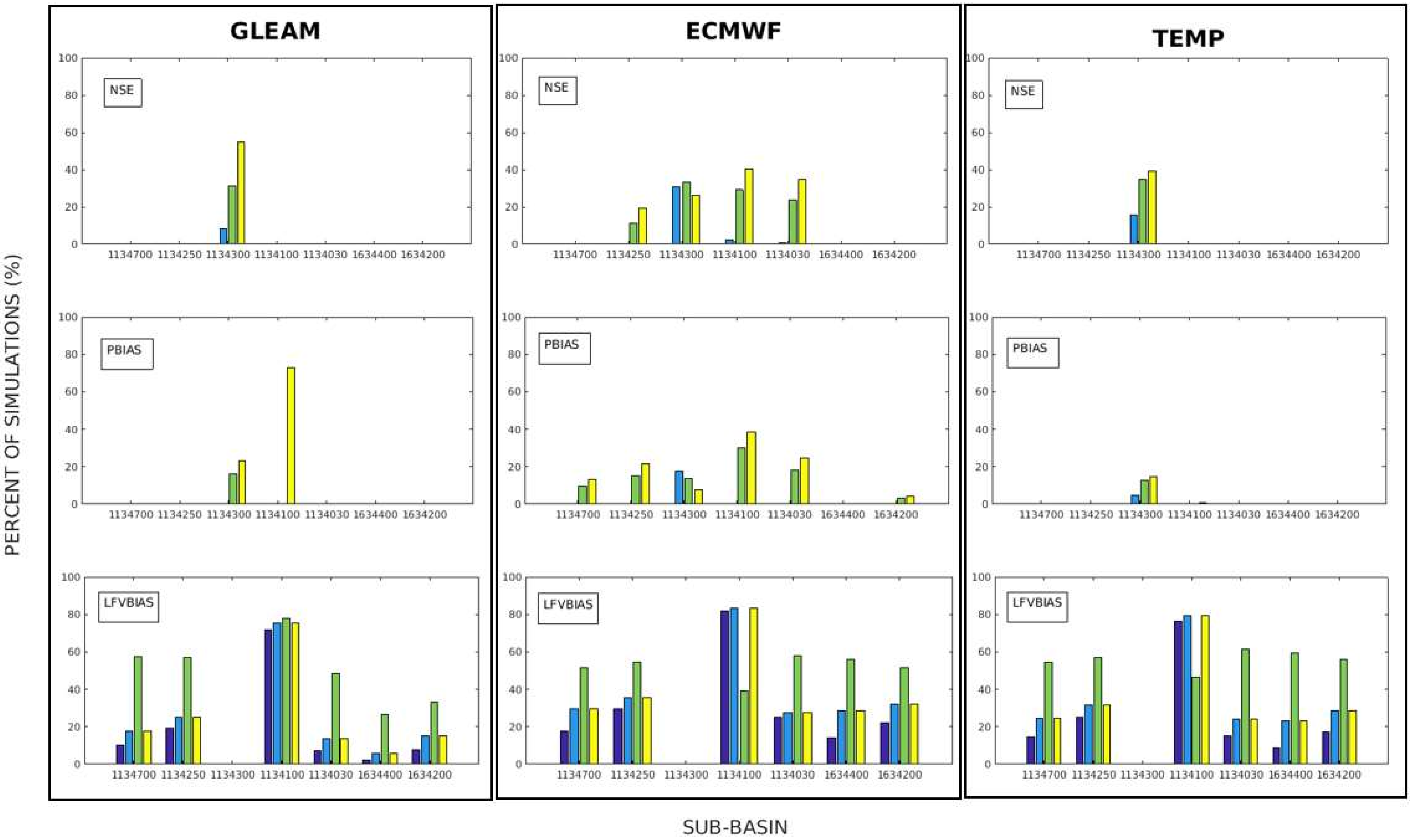


Figure 4.20. Percentage of simulations, for each model formulation for each model evaluation metric as calculated at each gauging station, that are over threshold values for 'acceptable' performance – NSE > 0, -20 < PBIAS < 20, -20 < LFVBIAS <20.

Table 4.13. Summary of the performance for the highest scoring simulation when calculated using three evaluation metrics, at each gauging station for each of the model formulations, using three PET datasets. Scores that have been highlighted are over the 'acceptable' performance threshold – NSE > 0, -20 < PBIAS < 20, -20 < LFVBIAS <20.

SUB-BASIN		**BENCHMARK**			**SRMAX INCREASED**			**SATEVAP ON**			**SATEVAP ON + SRMAX INCREASED**		
	PET USED	**GLEAM**	**ECMWF**	**TEMP**	**GLEAM**	**ECMWF**	**TEMP**	**GLEAM**	**ECMWF**	**TEMP**	**GLEAM**	**ECMWF**	**TEMP**
1134700	*NSE*	-50.62	-15.73	-42.16	-31.53	-3.83	-18.11	-0.29	-0.8	-9.0142	-23.48	-0.28	-5.7074
	PBIAS	447.56	231.76	391.95	332.08	92.996	227.03	267.9	0.0029	72.786	255.1	0.0059	60.309
	LFVBIAS	0.2688	-0.0263	-0.0195	-0.3059	0.0064	-0.0239	-0.0124	-0.0012	0.0006	0.0028	-0.0020	-0.0008
1134250	*NSE*	-10.74	-1.98	-8.2	-7.52	-0.06	-3.68	-7.38	0.53	-1.6508	-6.04	0.59	-0.7688
	PBIAS	456.35	148.11	264.32	259.33	68.86	180.84	223.7	0.010815	51.886	211.48	-0.0002	48.176
	LFVBIAS	0.6066	0.0024	0.0021	-0.0546	0.00181	-0.0076	-0.0015	-5.6902E-05	-0.0020	0.0039	-0.0016	0.0004
1134300	*NSE*	-1.62	-0.11	-1.59	0.47	0.49	0.53	0.63	0.59	0.61	0.64	0.59	0.61
	PBIAS	164.47	67.08	160.74	-0.09	-0.0058	-40.08	-0.03	-0.0300	-0.0149	0.04	0.0034	0.0069
	LFVBIAS	-	-	-	-	-	-	-	-	-	-	-	-
1134100	*NSE*	-5.97	-0.58	-4.34	-4.12	0.37	-1.68	-4.02	0.6	-0.5619	-3.19	0.62	-0.0116
	PBIAS	226.53	94.723	189.99	190.27	31.938	120.98	157.11	0.00578	21.28	0.01	-0.0049	12.319
	LFVBIAS	-0.0019	-0.0023	-0.0070	0.00731	-0.0051	0.0065	-0.0009	-0.0007	0.0045	0.0021	-0.0063	0.0008
1134030	*NSE*	-4.67	-0.49	-3.42	-3.87	0.23	-1.79	-3.78	0.53	-1.0254	-3.33	0.58	-0.5152
	PBIAS	228.19	202.42	190.84	206.03	56.22	142.11	185.21	-0.036527	53.532	181.62	-0.01231	47.862
	LFVBIAS	2.7201	0.0424	0.1781	1.7339	0.0046	-0.0107	0.0035	0.0026	0.0034	0.0045	-0.0001	0.0021
1634400	*NSE*	-13.98	-4.58	-10.87	-13.02	-2.95	-8.61	-13.39	-1.71	-7.3396	-12.8	-1.41	-6.6876
	PBIAS	369.06	202.42	315.63	354.82	149.06	263.11	342.5	47.05	180.26	340.33	39.872	171.21
	LFVBIAS	3.6572	-2.0942	0.1972	-0.2461	-0.0116	0.1058	0.2734	0.0013	0.0033	0.2789	0.0055	-0.0023
1634200	*NSE*	-4.59	-0.83	-3.02	-4.31	-0.44	-2.32	-4.46	-0.28	-2.3158	-4.26	-0.15	-1.9689
	PBIAS	203.19	92.569	158.44	194.33	60.694	127.16	187.97	8.7384	87.1	186.98	0.17507	76.976
	LFVBIAS	0.6691	-0.0731	0.2631	-0.6488	-0.0439	-0.0650	0.0063	0.0008	-0.0002	0.0010	0.0013	-0.0008

4.3.2. MULTI-OBJECTIVE EVALUATION

To find a 'behavioural' set of parameters and to evaluate model performance when combining all three metrics, these scores were individually ranked in order of best to worst performance. These ranks were then summed for each of the model simulations to get a combined rank, and then these ranks were sorted in order of lowest to highest. Each of the metrics were weighted with equal importance when calculating a combined score. In order to evaluate model performance with a multi-objective criteria, threshold values for each of the metrics were defined in which simulations needed to satisfy to be classed as 'behavioural'. These thresholds are based on ranges found in Moriasi et al. (2007) that are used to describe model performance as 'good'. These thresholds are NSE > 0.5, -10 < PBIAS <10, and -10 < LFVBIAS < 10. Uncertainty bounds for the 5^{th} and 95^{th} percentiles in the simulated flows were then calculated within a GLUE uncertainty analysis framework, and compared with observed flows in each of the sub-basins where a behavioural ensemble of parameter sets could be identified.

Behavioural simulations could only be identified in three sub-basins (1134030, 1134100, 1134250) using the model structure that includes evaporation from the saturated zone with the same parameterisation as the initial model simulations. Figure 4.21 is a comparison of the simulated flow timeseries for four sub-basins, split into upstream sub-basins where model performance is still not obtaining acceptable scores, and two downstream sub-basins where behavioural simulations could be identified. These simulations are from the saturated evaporation model structure with the same parameterisation as the initial model using ECMWF PET forcing data. It is clear that in the downstream sub-basins model performance has been drastically improved from the initial model implementation. All simulations seem to be able to capture the timing and the shape of the peaks in these sub-basins. Low flow volumes and the shape and timing of the recession flows is also captured well. In many years in the simulation period, the simulated peak flows are capturing the magnitude and volume of the high flows reasonably well, with a few exceptions of overprediction. In the upstream catchments, simulations during the dry season are capturing the low flows well. The timing and shape of the peaks also seem to capture the observed behaviour well, however, the magnitude of the flows is still not being accurately predicted. The simulated flows are also not receding at the same rate as is seen in the observed discharge data.

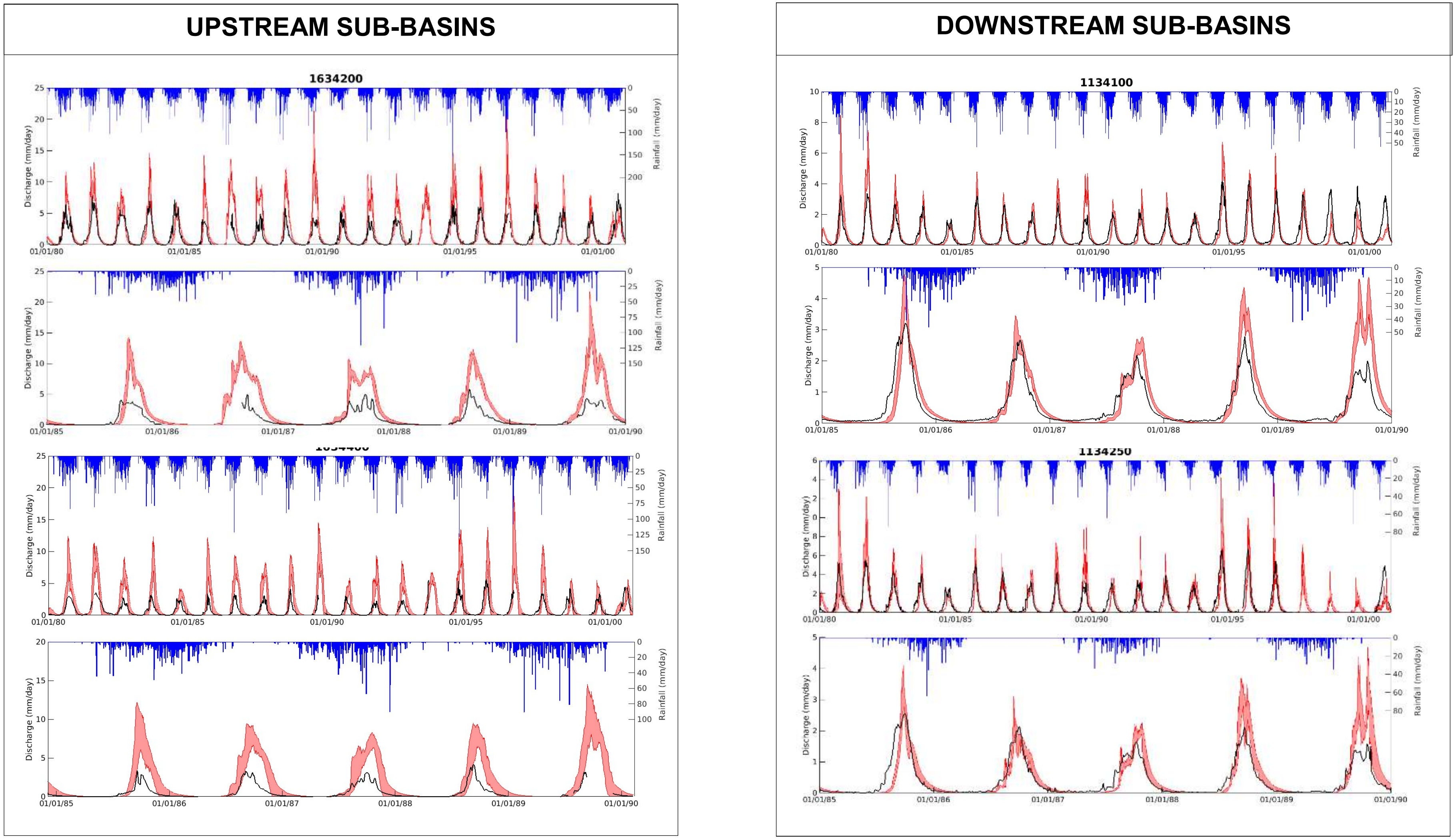


Figure 4.21. Uncertainty bounds for top 100 performing simulations in two upstream sub-basins, and behavioural simulations in two downstream sub-basins, in comparison with observed discharge at their downstream gauging station, using a modified model structure with evaporation taken from the saturated zone, with ECMWF PET forcing data.

Further analysis of the effect of modifying the model structure to incorporate evaporation from the saturated zone was carried out by looking more closely at the amount of evaporation that is taken from the soil root zone and the saturated zone at every timestep. Figure 4.22 shows this for one behavioural parameter set in the same two upstream and two downstream sub-basins. The amount of additional evaporation that is taken from the saturated zone varies between the sub-basins, with the highest rate occurring in the upstream sub-basins where it is almost doubled. As AET is at almost full potential rate during the wet season, no additional evaporation is taken from the saturated storage during these timesteps. However, during the dry season, where there is little to no rainfall input to the soil root zone, additional evaporation is being taken. Therefore, model performance is improved during the low flow periods of the simulations. Figure 4.12C also shows a comparison of the simulation with the same parameterisation but from the initial model structure, forced with ECMWF input data, for each of the sub-basins. This comparison clearly shows that the inclusion of evaporation from the saturated zone has improved overall model performance in the downstream sub-basins. In the upstream sub-basins, this comparison shows that there is little change in the magnitude of simulated flows using the two model structures, however, the saturated evaporation model is able to improve the timing of the peak flows.

Although behavioural simulations, as defined by thresholds of ‘good’ performance (Moriasi et al., 2007), could only be identified for three sub-basins, it should be noted that model performance in all sub-basins has been drastically improved that have been evaluated in study when using the modified model structure including saturated zone evaporation. This is summarised by the simulation performance scores given in Table 4.13. This indicates the importance of the development of flexible model structures, especially at large scales, where a ‘traditional’ fixed model structure would struggle to represent the spatial heterogeneities. It also reflects the importance of being able to test different hypotheses of hydrological behaviour in large and data sparse catchments where we do not have observational data and field measurements to inform our understanding of basin functioning.

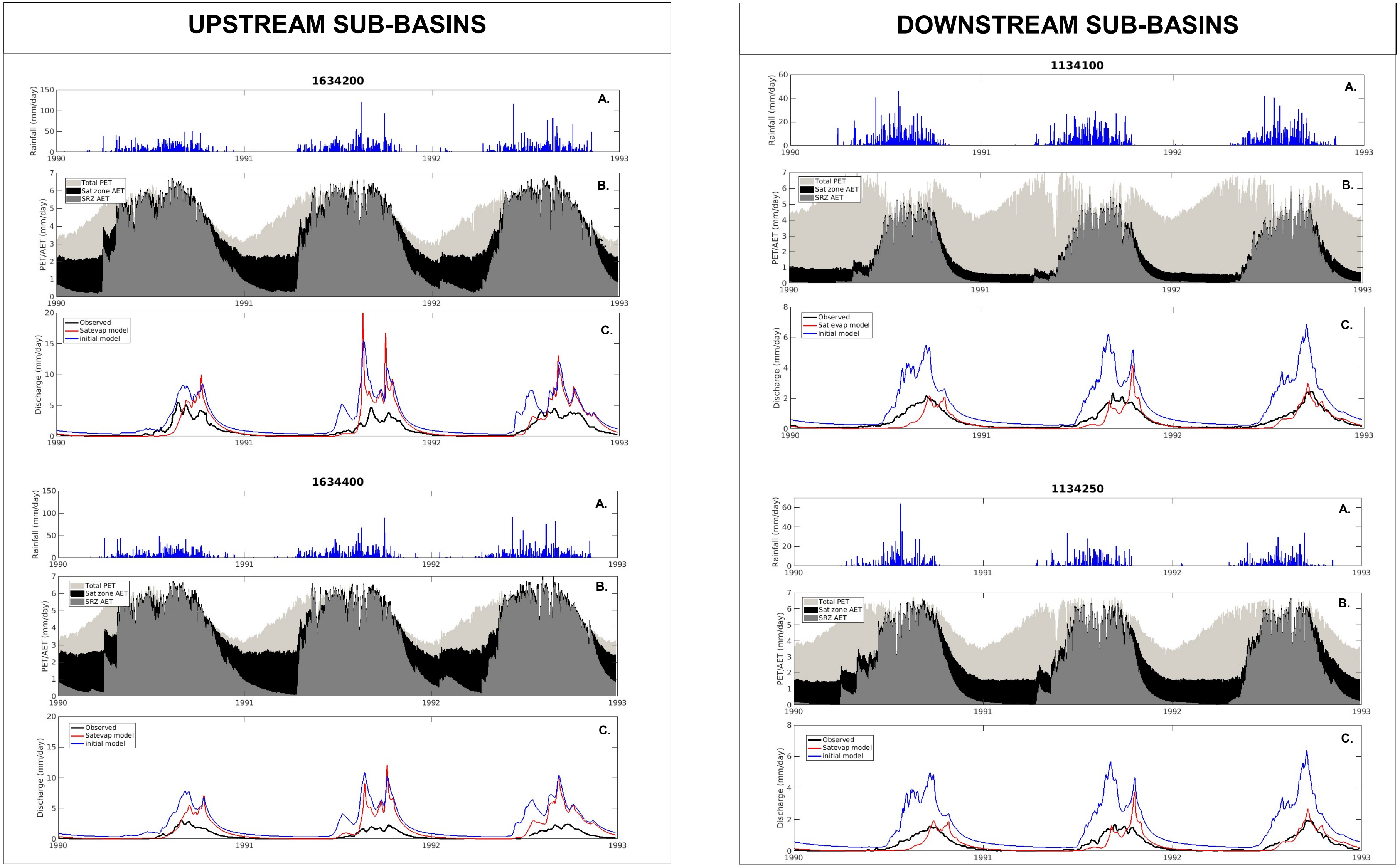


Figure 4.22. For one behavioural parameter set from satevap model simulations forced with ECMWF PET data, for three years of the simulation period, in two upstream sub-basins and two downstream sub-basins, as defined by their downstream gauging station: A. Mean daily precipitation, B. Comparison of input PET, AET from the saturated zone, and AET from the soil root zone, 3. Comparison of simulated flow from the initial model and satevap model with the observed discharge at each gauge.

The ability of model simulations to capture the inter- and intra-annual variability and cycles of the observed discharge can be characterised by different properties (Gudmundsson et al., 2012). The mean annual value is a measure of the water balance in a catchment and can give insight into the long-term water balance. The difference in the highest to the lowest simulated flow value, or the amplitude, is a measure of how pronounced the seasonal cycle of model simulations is when compared to the observed flow.

Figure 4.23 summaries the ability of the behavioral ensemble in three sub-basins to capture this inter-annual and seasonal variability. Simulations in sub-basin 1134100 manage to simulate the long-term mean annual cycle of the observed discharge very well, with a median relative bias of approximately 0.005. There is also the smallest range in this sub-basin too. However, simulations ability to capture the amplitude of the mean annual cycle is not of a similar good performance, and sub-basins 1134030 and 1134250 are much better at simulating this. The model's ability to simulate the mean monthly variability of the observed discharge is good for each of the sub-basins. For many months, the median of the ensemble of simulations is close to zero, and variation around this median is small. The ranges for the months that immediately follow the wet season (i.e. November, December and January) are larger in comparison with other months, especially for January in 1134100 and 1134250. This may be because these particular gauging stations are downstream from the two major dams on the Upper Niger river, and river flows are controlled during the dry season, and so simulated flows may not reflect observations as well during these months.

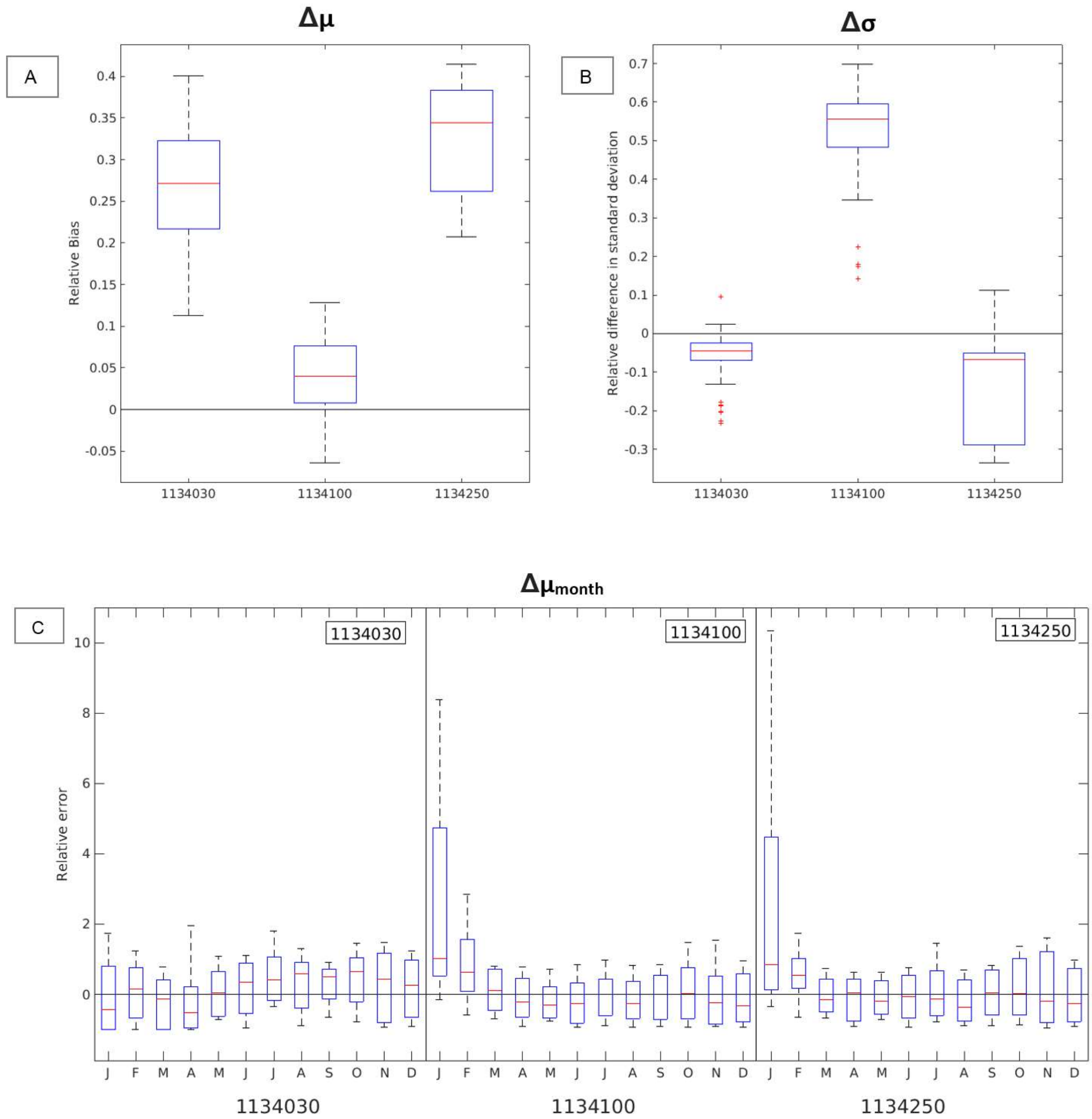


Figure 4.23. Evaluation of 'behavioural' simulations to capture the annual and monthly seasonal cycle of observed flow characterised by: A. relative bias in the mean annual runoff, B. relative difference in standard deviation of observed and simulated flow, and C. relative error in mean for individual months.

5. DISCUSSION

From the literature discussed in Section 2, and the results that presented in Section 4, the following research questions (that were defined in Section 1.4) will now be discussed:

1. Evaluate the performance of DECIPHeR in a large and data scarce model domain, using the Upper Niger basin as a case study location.
2. Investigate the different sources of uncertainty and analyse the impacts that these have on the hydrological models performance.
3. Highlight areas of this new framework where future improvements and developments are required for modelling at large scales.

5.1. MODEL PERFORMANCE EVALUATION

DECIPHeR has been set up for the Upper Niger basin and run with three different PET input datasets to explore uncertainty. 10,000 randomly sampled parameter sets were used for a 21-year simulation period, 01/01/1980 – 31/12/2000. Three evaluation metrics, each testing model capability to reproduce different aspects of hydrological behaviour, were calculated for simulations at seven gauging stations across the basin. Parameterisation and model structure were adjusted and modified to incorporate and better represent the processes that are occurring in the catchment. The results presented in Section 4 show that with the modified model structure that incorporates evaporation being taken from the saturated zone, model performance is significantly improved across the basin. However, model performance was ‘good’, determined by threshold values in the performance measures, in only four-sub-basins.

The headwater catchments of the Upper Niger are in the humid and tropical climate zone of the basin. This region is characterised by large volumes of rainfall during the wet season from March-August. This is where the Niger gains the majority of its discharge that is then propagated downstream. However, sub-basins 1634200 and 1634400 that are in this zone have very low runoff coefficients, of 0.21 and 0.12 respectively. It is a common finding in hydrological modelling studies that model performance suffers in catchments with low runoff coefficients, and was found in the initial set up and evaluation of DECIPHeR in the UK in Coxon et al. (2018). Catchments with low-runoff coefficients (i.e. less the 0.2) are often found in regions that are dominated by groundwater influences. An investigation into the groundwater properties of the headwater sub-basins was carried out. From global maps produced by the British Geological Survey (BGS), it was found that groundwater stores in the headwaters of the Niger were described as shallow and local aquifers with fractured connectivity, with low recharge potential, and therefore it is assumed that this is not a dominant process that is not being represented by DECIPHeR for these sub-basins.

The model structure of DECIPHeR currently maintains mass balance, and therefore does not allow water to be lost from the basin. This combined with the model dynamics may also be a reason for why model performance is poorer in the headwater catchments in comparison with the downstream, drier sub-basins. In these parts of the Upper Niger Basin, daily precipitation estimates are far greater than that of PET. Therefore, there is a much greater volume of input rainfall on the daily timestep in the tropical wet headwaters than can be removed by evaporation, even when being taken at the full potential rate. This is still the case when evaporation from the saturated zone was included in the model, as daily PET inputs during the wet season and where peak flows occur were relatively too low, and therefore could not reduce runoff enough to accurately simulate observed discharge.

5.2. UNCERTAINTY ANALYSIS

5.2.1. MODEL STRUCTURAL UNCERTAINTY

From the results presented in Section 4, it is clear that by changing the model structure for the Upper Niger basin, model performance has been improved when evaluated using the metrics in this study in all sub-basins. However, it is only for sub-basins 1134030, 1134100, 1134300, and 1134250 where ‘good’ model performance can be achieved. Sub-basins 1634200 and 1634400 are the natural headwater catchments for the Upper Niger basin (i.e. minimal impacts from anthropogenic influences), and therefore it is surprising that good model performance has not been able to be achieved in these catchments. A common finding in hydrological modelling studies is that better performance is found in wetter catchments, and generally poorer performance is found in dry catchments (Gosling and Arnell, 2011; Newman et al., 2015; McMillian et al., 2016). However, the opposite has occurred here in this study, where the dryer downstream catchments have achieved much higher model performance. There is a clear seasonal water balance issue in the headwater catchments. This is caused by a combination of the discrepancies and errors in the input data, and a lack of knowledge of the catchment functioning to inform our decisions of how best to represent the processes basin with the model structure. This result may also be because of the complex topography that is found in the headwater catchments, as they are in the Guinea Highlands, and there is relatively little topographic change for much of the Upper Niger basin once the river channel leaves Guinea and flows through Mali. There is only approximately 60m of elevation change between gauging station 1134030 and 1134250, which are around 700km apart. However, this would require further investigation of the effect of the representation of topographic complexity in these sub-basins on the simulations of river flows.

In many previous hydrological modelling studies in the Upper Niger basin authors have concluded that the reason for overpredicition of simulated discharge is due to an underestimation in, or a misrepresentation of, evaporation (Dezetter et al., 2008; Dadson et al., 2010; Huang et al., 2017; Andersson et al., 2017a). Andersson et al. (2017a,b) and Thompson et al. (2016, 2017) found that by changing the hydrological model structure used, model simulations performance could be improved. In Andersson et al. (2017a), the HYPE model concept (Hydrological Predictions for the Environment, originally developed for use in Sweden) was applied to the Niger Basin. The results of the study showed that the model was able to capture the annual cycle of the river flows, but it was unable to simulate the magnitudes, gaining an average NSE score of -1 across the 56 sub-basins. Evaporation was one of the processes that was identified as needing an improved description in the model code to be able to improve the simulations produced. The Hargreaves method of estimating PET was implemented in the new version of the model for the Niger, as this this was found by Oudin et al. (2005) to be more suitable in dry catchments where evaporation is dominant in the water balance. With the process refinements in the Niger-HYPE 2.0 model, performance was found to be significantly better, with an average NSE score of 0.4 across the basin. In Thompson et al. (2016, 2017) a semi-distributed hydrological model of the Upper Niger was used to investigate the effect of climate change projections on future river flows. In this study an out of bank flood inundation element was added to represent the influences of the Inner Niger delta on river discharges in the basin. The model also incorporated modules to represent the two large dams on the Upper Niger, the Selingue Dam and the Markala Barrage, and this was based on previous analysis of the river flow data by Zwarts et al. (2005). There were eleven sub-basins in this study's domain, and when evaluated with NSE, river discharge simulations were classified as 'excellent' for the calibration period, and 'very good' during the evaluation period. These studies indicate that in order to successfully apply hydrological models in the Upper Niger basin, processes that are represented need refinement.

5.2.2. INPUT DATA ERRORS AND UNCERTAINTY

POTENTIAL EVAPOTRANSPIRATION DATA

This study has highlighted the large uncertainty in global evapotranspiration (ET) datasets. ET is one of the most important fluxes in the hydrological cycle, and therefore accurate representations are crucial for making predictions of streamflow in catchments. Generally, ET has been the most difficult meteorological aspect of the water balance to estimate (Lettenmaier and Famigletti, 2006). The uncertainty and inconsistencies between remote-sensing driven global products is due to the combination of input datasets that are used to estimate global ET using algorithms (e.g. Priestly-Taylor, Pennman-Monteith, etc). These data include net radiation, meteorology (e.g. water vapour pressure, air temperate, and wind speeds), and vegetation (e.g. land cover, soil type, soil moisture content) (Fisher et al., 2008; Mu et al., 2007; 2011). Also, direct ground measurements of ET are not possible on such large scales, and are only available in small regions using methods such as flux tower, but these are limited to a small number of sites mostly in developed countries (Miralles et al., 2016). This makes evaluation of global PET datasets outside of North America and Europe extremely difficult (Trambauer et al., 2014).

Due to the uncertainty and inconsistencies in PET input data, it is common in hydrological modelling studies for this process to be simplified in catchments, for example assuming that PET is the same in each month every year (Samain and Pauwels, 2013). There is debate in the literature about how best to represent PET in hydrological models in order to make accurate predictions of discharge. For example, Oudin et al. (2005) used four hydrological models in 308 catchments in France, Australia and the United States that covered a large range of climatic properties, and concluded that daily observations of PET are not necessary for streamflow simulation accuracy. A similar result was found by Dezetter et al. (2008). Their objective was to assess whether there is a single data-model combination that can be used in West Africa for the simulation of rainfall-runoff. Three PET input grids were used, but the two hydrological models used in the study (GR2M and Water Balance Model) were found to not be sensitive to these. The results found in this study, however, contrast with this finding. Here, it was found that using three spatially varying global PET datasets drastically changed model performance. The PET grids used by Dezetter et al. (2008) were all derived from the Penman formula, and used monthly estimates with spatial resolution 0.5°x0.5°. The differences between each of the grids was found to be around 10%. However, the PET datasets used here in this study have been derived from different algorithms and equations, provide daily estimates, and the discrepancies between (and within) each of the products has been shown to be large.

RAINFALL DATA

Despite the clear uncertainty that is inherent in these large global datasets, only one rainfall dataset was used in this study. MSWEP global precipitation data has been evaluated in the literature (Dembele and Zwart, 2016; Beck et al., 2017; Shalou et al., 2017) and has been applied in hydrological modelling studies in data sparse catchments (Alijanian et al., 2017; Nair and Indu, 2017; Shalou et al., 2017), where simulations have obtained good performance. Also, estimates in the dataset are similar to those that are reported in the literature, which are based on both gauge measurements and from other satellite based observations of precipitation (e.g. Zwarts et al., 2005; Zwarts, 2011; Thompson et al., 2017). However, there are inevitably large errors and uncertainty in a gridded rainfall dataset. Precipitation patterns have an ‘intrinsic irregularity’ (Molini et al., 2001) which makes it difficult to measure the real time and space evolution of precipitation fields. Therefore, accounting for this spatial and temporal variability in a gridded global dataset is extremely difficult. The accurate representation of rainfall is needed in hydrological modelling in order to produce the most physically realistic simulations (Beven, 2004). The impacts of rainfall errors on predicted flow has been widely discussed in the literature (e.g. Sun et al., 2008; Bardossy and Das, 2008; Moulin et al., 2009; McMillian et al., 2011). As mentioned above, the resolution of the input rainfall data is relatively coarse, at approximately $10km^2$, and therefore this is likely underpredicting much of the spatial heterogeneity across the basin.

DISCHARGE DATA

Another major source of uncertainty when evaluating the performance of hydrological models in the discharge observational data. This is often the only data available to calibrate and validate hydrological models at the catchment scale (Chang et al., 2017), however it is recognised as one of the key sources of error in modelling studies (Dottori et al., 2009; Westerberg et al., 2011; McMillan et al., 2012; Coxon et al., 2015). The discharge data used for model evaluation in this study was taken from the GRDC. The quality of the discharge data in Western Africa in this database has been declining since the 1980s when many hydrometric stations were closed and reduced to a minimum (Nkamdjou and Bedimo, 2008). The data that is available is incomplete and of questionable quality. This is likely affecting the ability to evaluate model performance, as this is the only reference to the hydrological behaviour in the Upper Niger basin. The densest network of gauging data occurred in the 1950/60s, but as there was no global precipitation and PET datasets that overlapped with this period, thus, it was not used as part of the simulation period for this model evaluation study.

5.3. STUDY LIMITATIONS AND FUTURE MODEL DEVELOPMENTS

PROCESS REPRESENTATION AND REFINEMENT

One of the main limitations of this study is that the saturated zone evaporation model structure was implemented across the whole basin. There has been no experimentation of using different model structures in different parts of the landscape in this study, and it is recognised that this likely to be of importance in a basin as large as the Upper Niger, with such varying hydro-climatic conditions. At the downstream region of the model domain is the Inner Niger Delta. This is a seasonal floodplain that can be as large as 40,000km^2 during a high flood year. This is a large amount of water to be on the surface of the basin and reduced the discharge between its upstream and downstream gauge by approximately 20% (Schuol and Abbaspour, 2006). Therefore, a flood inundation module could be added to this part of the basin in future improvements to this model set-up. An area-discharge relationship between the discharge entering the upstream boundary of the sub-basin and the flood extent on the land surface could be calculated and implemented in future model structural changes in the DECIPHeR Niger model.

As mentioned above, the headwater sub-basins in this model domain have more complex topography than the downstream catchments. In future refinements to the DECIPHeR Niger model, experimentation of the spatial complexity in these sub-basins could be investigated. For example, during the DTA when hydrologic similarity is defined in order to classify the HRUs, additional classes of accumulated area and slope could be added in parts of the landscape where there is more complex topographic features, such as the Guinean Highlands located in the headwaters of the Upper Niger. This would result in more HRUs being defined in this part of the landscape, and this may lead to a better representation of the rainfall-runoff relationship locally to these areas.

ANTHROPOGENIC INFLUENCES

Human impacts have a profound impact on river discharges and water related hazards, such as flooding and droughts (Padowski et al., 2015; van Loon et al., 2016; Liu et al., 2017), and there have been research efforts to parameterise human activity in hydrological models (Bierkens, 2015; Pokhrel et al., 2016). These parameterisations aim to include representation of dams and reservoirs, and changes in land use, land cover, and land management (Voisin et al., 2013; Pokhrel et al., 2016; Wada et al., 2016, 2017). However, there are large differences in simulated discharge using models which have incorporated human influence parameterisations, due to the errors and uncertainty in the data used in the calibration of these parameters, etc (Duan et al., 2006; Doll et al., 2016).

Human activity and impacts on the Niger River are varied across the basin. Annual discharges have been decreasing in the Niger River in the last few decades (Zwarts, 2011), and this has been accounted for by changes in rainfall in the basin. However, Ferry et al. (2011) says it is likely to be a combination of many factors including impacts of large dams, and usually assumed negligible, small-scale water abstractions. There are many small-scale abstractions and diversions from the river for irrigation and agricultural used. However, there is a lack of data for these small schemes, and therefore incorporating them into the model structure is difficult. There are also two large dams in the Upper Niger basin that were in operation during the simulation time period used in this study, the Selingue Dam and the Markala barrage. The Selingue dam is used for storing water for hydroelectricity power production and irrigation schemes. Zwarts (2011) suggests that this dam has a large effect on the hydrology of the Niger. The reservoir created by the dam leads to a water loss from evaporation in the region. It also effects the seasonal river discharges. A comparison of the inflow to the reservoir before the dam, which represent natural river flow, and the outflow that is controlled, shows that the discharge is reduced by 61% in august and 36% in September (Zwarts et al., 2005). The Markala barrage is used for irrigation managed by Office du Niger. Water abstraction and diversions from the Niger for this project were at the same level between 1988 and 2009 (Zwarts, 2011). This affected flow depending on whether it was a high or a low river flow year. In high flow years, only 7% of the discharge was diverted, however this increased to 16% in low flow years (Zwarts et al., 2005).

It is clear that these abstractions and diversions in the Upper Niger basin have a profound effect on the discharge, and with the current model formulation this is not being accounted for. In future improvement and developments of the DECIPHeR Niger model, these human impacts need to be accounted for in the structure. Due to the flexibility of the modelling framework used by DECIPHeR, changes to the model structure can be made in different parts of the model domain, and therefore these processes could be added locally to where the effects are occurring. However, implementing structural modification of this nature are constrained by the data available.

6. CONCLUSIONS

6.1. AIMS OF DISSERTATION

This modelling study had three core aims and research questions.

1. Evaluate the performance of a new flexible modelling framework, DECIPHeR, in a large and data scarce model domain, using the Upper Niger basin as a case study.

DECIPHeR (Dynamic fluxEs and ConnectIvity for Predictions of HydRology, Coxon et al., 2018) is a new flexible hydrological modelling framework for uncertain flow simulation and prediction at catchment to continental scales. The model can be modified and adapted to suit specific hydrological settings and available data for the region of interest. DECIPHeR has so far only been applied to the UK, as an initial demonstration of use in a large-scale application and benchmark for model performance. A single simple model structure was used across the model domain. However, the UK is a well gauged, data rich location, and therefore further model evaluation studies are required, particularly in large and data scarce locations. Therefore, in this study, DECIPHeR has been applied in the Upper Niger basin. This was chosen as an appropriate study site for a number of reasons: 1) There are very few ground observations of catchment functioning to inform model structural choices, 2) the data that is used is from open-source global products, and this study provides an evaluation of these datasets in hydrological modelling, and 3) there is a strong gradient in hydroclimatic variability from the upstream headwaters to the downstream outlet, which provides a test of the rainfall-runoff model's capacity to represent these different dynamics.

The initial model structure described in Coxon et al. (2018) was applied across the Upper Niger basin as a benchmark of model performance. When evaluated with three performance metrics (NSE, PBIAS and LFVBIAS), simulations were poor in all sub-basins, with a significant overestimation of peak flow magnitudes, resulting from a combination of input data errors, and structural insufficiency. It was concluded from these initial results that some key processes were not being represented with this model set-up. An investigation of the groundwater influence in the basin was carried out, to assess whether this was a dominant controlling factor in the Upper Niger. Aquifers in the region are described as shallow and local with fractured connectivity with low recharge potential by the British Geological Survey (BGS), and therefore is not a dominant characteristic in this basin.

In previous studies in the Upper Niger basin, it has been found that evaporation is a dominant process occurring, and where it's representation within the model structure is modified, model simulations are significantly improved (Thompson et al., 2016, 2017; Andersson et al., 2017a,b). Therefore, the rainfall-runoff model structure in DECIPHeR was modified to change the representation of evaporation for the Upper Niger basin. A saturated zone evaporation module was added to the model code. This was model structure was then applied across the model domain, and model performance was significantly improved in all sub-basins. It was with this model structure that when all three evaluation criteria were combined, behavioural simulations could be identified, as defined by thresholds in each of the performance metrics.

2. Investigate the different sources of uncertainty and analyse the impacts that these have on the hydrological model's performance.

There are three major sources of uncertainty in hydrological modelling studies – input and output data, parametric and structural. Each of these sources has been investigated in this study and model simulations have shown to have different levels of sensitivity to these uncertainties.

The input data that was available for the Upper Niger basin was taken from open-source global datasets. One rainfall product (MSWEP) and three PET products (GLEAM, ECMWF, estimates based on relationship between temperature and latitude) were used, each with 0.5°x0.5° resolution. Only one rainfall product was used as the estimates in this dataset agreed well with what is reported in the literature. However, three different PET products were used as there were large inconsistencies between each of the PET datasets. This is expected due to the many available methods of estimating PET, and the different input data requirements of these algorithms. Simulated flow was found to be very sensitive to the errors in these PET datasets, with significant gains for all performance metrics when the model was forced with ECMWF PET. This highlights the importance of evaluating different global datasets when used for hydrological modelling, as they can have a profound effect on results.

To explore parametric uncertainty, model simulations were run within a Monte Carlo framework, where 10,000 sets were randomly sampled for each parameter between a user defined lower and upper boundary, which have an assumed uniform distribution. For initial model simulations, parameter bounds were set as the same as those that are used in the application of DECIPHeR in the UK (Coxon et al., 2018), as these were set to be wide to cover a large range of hydro-climatic conditions in a national application. After evaluation of these initial results, the upper boundary for parameter SRmax was then increased. This was in response to model simulations overestimating flow magnitudes in all sub-basins in the model domain. SRmax is the parameter that controls the soil root zone, which is where evaporation is taken directly from, and is the only way water can be lost from the system using the initial model structure described in Coxon et al. (2018). By doing this, model performance in all sub-basins was improved.

To investigate structural uncertainty, the model structure used in Coxon et al. (2018) was used as a benchmark for model performance. This structure was then modified to include an additional module, where any residual evaporative demand that was not met by the soil root zone could be taken from the saturated zone and then implemented across the whole model domain. By including this new module to the rainfall-runoff model performance was significantly improved. This highlights the importance of hypothesis testing in hydrological modelling studies of large and data scarce river basins where little information about catchment functioning is available to inform a priori structure choices.

3. Highlight areas of this new framework where future improvements and developments are required for modelling at large scales.

This evaluation study of DECIPHeR has showed this framework's ability to be successfully set up and run in large and data scarce river basins. However, 'good' model performance and behavioural simulations could only be identified in four of the seven sub-basins. This indicates that there are still parts of the basin that are controlled by some dominant hydrological features that have not been represented either by the original model structure applied in the initial benchmark simulations or in the modified structure that includes evaporation from the saturated zone. An example of one of these features is the Inner Niger Delta. This is a large seasonal inland floodplain, which can be as large as 40,000km^2. This delta increases evaporation locally due to the large volume of open water on the earth's surface. It is reported to reduce the discharge from its upstream gauge to its downstream gauge by around 20-30%. This is currently not being represented in the model, and in an improved future DECIPHeR Niger model, a seasonal flood inundation module could be added to this part of the landscape to account for this. This could then be transferred to other model set ups in other river basins that also have hydrological features similar to this delta.

6.2. FUTURE MODEL IMPROVEMENTS AND DEVELOPMENT

There are a number of limitations associated with this investigation. Firstly, only one model structural change was made for the Upper Niger basin, and this was applied homogenously across the domain. As has been discussed throughout this dissertation, there are large differences in the dominant hydrological processes occurring across the basin, and it is likely that one structure implemented across the whole domain can account for these large differences. In future improvements to the DECIPHeR Niger model, experimentation and hypotheses testing of catchment functioning represented with the model structure is needed. The parameters used in this study were also applied homogenously across the whole model domain. Although upper boundaries for parameter SRmax were increased to allow for more evaporation to be directly taken from the soil root zone, there was no experimentation of the impacts that the parameters have on simulation river flows in the Upper

Niger basin. Ongoing work in the development of the DECIPHeR framework aims to implement the multiscale parameter regionalisation (MPR) technique, which is a method of linking model parameters to geophysical catchment characteristics through transfer functions applied at the finest possible resolution (Samaniego et al., 2010), i.e. HRUs defined as the size of one grid cell. The coefficients of the transfer functions are then calibrated, and the parameters are upscaled to produce spatially consistent fields of model parameters for any spatial resolution across an entire model domain. The purpose of this is to produce parameter values that are not dependent on scale, and can be applied in catchments where physical properties can be derived, but the functioning of the catchment is unknown.

Another limitation that is identified in this study is that, although the uncertainty associated with input data errors was explored, only one rainfall product and three PET products were used. This is a small sample, and additional global data products would need to be used as input data to investigate the impact that input data uncertainty has on model simulations more thoroughly. The large differences between the three different PET datasets used in this study highlights the large uncertainty that is inherent in global data products. These datasets are the only source of hydrological input data in many of the world's river basins, and so a rigorous evaluation of their influence on model simulations is needed in order to produce robust predictions of flow in large-scale modelling studies. Also, in this study, only spatially varying climatic inputs and topographic data were used to define similarity within the landscape from which HRUs were then classified. DECIPHeR gives the user the option of including addition datasets that can be used to discretise the model domain, such as soil types, geology, land use and land cover, etc. However, these datasets are often not available for many of the world's large river basins. This limits the framework's ability to group together areas in the landscape that have similar characteristics, and therefore limits the level of complexity of the hydrological processes that are represented in the model structure.

6.3. CONCLUDING REMARKS

This study has shown the capacity of DECIPHeR to be applied in large and data scarce domains. Although simulations in only four of the seven sub-basins were able to achieve 'good' performance when evaluated with three metrics, predictions of flow were improved significantly in all sub-basins when the parameterisation and model structure were adjusted to better represent the Upper Niger. This is the first application of DECIPHeR outside of the UK – a data rich and well gauged location – and this study has proved that this flexible modelling framework is a good tool for testing different hypotheses of basin functioning where data is not available to inform decisions. However, additional work is needed in the future for further developments and improvements to this framework.

7. REFERNCES

Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R. (2007) Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT, *Journal of Hydrology*, 333, 413-430

Abbott, M.B., Bathrust, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J. (1986) An introduction to European hydrological system – systeme hydrologique Europeen (SHE) Part 2. Structure of a physically based distributed modelling system, *Journal of Hydrology*, 87, 61-77

Abdelnour, A., Stieglitz, M., Pan, F., McKane, R. (2011) Catchment hydrological responses to forest harvest amount and spatial pattern, *Water Resources Research*, 47, W09521

Aich, V., Liersch, S., Vetter, T., Andersson, J., Muller, E., Hattermann, F. (2015) Climate or land use? – attribution of changes in river flooding in the Sahel Zone, *Wate*r, 7, 2796-2820

Aich, V., Liersch, T., Vetter, T., Fournet, S., Andersson, J., Calmanti, S., van Weert, F.H.A., Paton, E.N. (2016) Flood projections within the Niger River Basin under future land use and climate change, *Science of the Total Environment*, 526, 666-677

Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S. (2004) Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, *Journal of Hydrology*, 298(1-4), 112-135

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T. and Siebert, S. (2003) Global estimation of water withdrawals and availability under current and "business as usual" conditions, *Hydrological Sciences Journal.*, 48(3), 339–348

Alijanian, M., Rakhshandehroo, G. R., Mishra, A. K., Dehghani, M. (2017) Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran, *International Journal of Climatology*, 37, 4896–4914

Allen, R.G., Tasumi, M., Morse, A., Trezza, R., Weight, J.L., Bastiaanssen, W., Kramber, W., Lorite, I., Robison, C.W. (2007) Satellite-based energy balance for mapping evapotranspiration with internalised calibration (METRIC) - applications, *Journal of Irrigation and Drainage Engineering*, 133(4), 395-406

Amadou, A., Djibo Abdouramane, G., Ousmane, S., Sanda Ibrah, S. and Sittichok, K. (2014) Changes to flow regime on the Niger River at Koulikoro under a changing climate, *Hydrological Sciences Journal*, 60(10), 1709–1723

Andersson, J., Arheimer, B., Traore, F., Gustafsson, D., Ali, A. (2017a) Process refinements improve a hydrological model concept applied to the Niger River basin, Hydrological Processes, 31(25), 4540-4554

Andersson, J., Ali, A., Arheimer, B., Gustafsson, D., Minougou, B. (2017b) Proving peak river flow statistics and forecasting in the Niger River basin, *Physics and Chemistry of the Earth, 100*, 3-12

Andreassian, V., Le Moine, N., Perrin, C., Ramons,M., Oudin, L., Mathevet, T., Lerat, J., Berthet, L. (2012) All that glitters is not gold: the case of calibrating hydrological models, *Hydrological Processes*, 26, 2206-2210

Apel, H., Merz, B., Thieken, A.H. (2008) Quantification of uncertainties in flood risk assessments, *International Journal of River Basin Management*, 6(2), 149-162

Apip, Sayama, T., Tachikawa, Y., Takara, K. (2012) Spatial lumping of a distributed rainfall-sediment-runoff model and its effective lumping scale, *Hydrological Processes*, 26, 855–871

Archfield, S.A., Clark, M., Arheimer, B., Hay, L.E., Farmer, W.H., McMillan, H., Seibert, J., Kiang, J.E., Wagener, T., Bock, A., Hakala, K., Andreassian, V., Attinger, S., Viglione, A., Knight, R.R., Over, T.M. (2015) Accelerating advances in continental domain hydrologic modelling, *Water Resources Research*, 51, 10078-10091

Arheimer, B., Dahne, J., Donnelly, C., Lindstrom, G., Stromqvist, J. (2012) Water and nutrient simulations using the HYPE model for Sweden vs. the Baltic Sea basin – influence of input-data quality and scale, *Hydrology Research*, 43(4), 315-329

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., Javelle, P. (2011) Sensitivity of hydrological models to uncertainty in input, *Hydrological Sciences Journal*, 56(3), 397-410

Arnell, N.W. (1999) The effect of climate change on hydrological regimes in Europe: a continental perspective, *Global Environmental change*, 9(1), 5-23

Ashouri, H., Hsu, K., Sorooshian, S., Braithwaite, D. (2015) PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies, *Bulletin of American Meteorological Society*, 96(1), 69-83

Badgley, G., Fisher, J.B., Jimenez, C., Tu, K.P., Vinukollu, R. (2015) On uncertainty in global terrestrial evapotranspiration estimates from choice of input forcing datasets, Journal of Hydrometeorology, 16, 1449-1455

Bai, Y., T. Wagener, and P. Reed (2009), A top-down framework for watershed model evaluation and selection under uncertainty, *Environmental Modelling & Software*, 24(8), 901-916

Baldocchi, D.D., Ryu, Y. (2011) A synthesis of forest evaporation fluxes – from days to years as measured with eddy covariance, *Ecological Studies: Analysis and Synthesis*, 216, 101-116

Balsamo. G., Viterbo, P., Beljaars, A., van den Hurk, B., Hirschi, M., Betts, A.K., Scipal, K. (2009) A revised hydrology for the ECMWF Model: Verification from field site to terrestrial water storage and impact in the integrated forecast system, *Journal of Hydrometeorology*, 10(3),623-643

Bárdossy, A., Das, T. (2008)Influence of rainfall observation network on model calibration and application, *Hydrology and Earth System Science*,12, 77–89

Bates, P. D., Pappenbergen, F. and Romanowicz, R. (2013) Uncertainty in flood inundation modelling. In: K. Beven and J. Hall (eds), Applied Uncertainty Analysis for Flood Risk Management. Imperial College Press: London. 230-261

Beck, H. et al. 92017) Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modelling, *Hydrology and Earth System Sciences*, 21, 6201-6217

Beck, C., Grieser, J.,Rudolf, B. (2005) A new monthly precipitation climatology for the global land areas for the period 1951 to 2000. DWD Climate Status Rep. 2004, 181–190

Beck, H.E., van Dijk, A., de Roo, A., Dutra, E., Fink, G., Orth, R., Schellekens, J. (2016) Global evaluation of runoff from ten state-of-the-art hydrological models, *Hydrology and Earth Systems Science*, 21(6), 2881-2903

Beck, H., Wood, E., Pan, M., Fisher, C., Miralles, D., van Dijk, A., McVicar, T., Adler, R. (2018) MSWEP V2 global 3-hourly 0.1 degree precipitation: methodology and quantitative assessment, *Bulletin of the American Meteorological Society*

Bergström, S. (1992) The HBV model - its structure and applications, SMHI Hydrology, RH No.4, Norrköping, 35 pp

Best, M.J. et al. (2011) The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, *Geoscientific Model Development*, 4, 677-699

Beven, K.J. (2000) Uniqueness of place and process representations in hydrological modelling, *Hydrology and Earth Systems Sciences*, 4, 203-213

Beven, K.J. (2001) On hypothesis testing in hydrology, *Hydrological Processes*, 15(9), 1655-1657

Beven, K. (2009) *Environmental modelling: an uncertain future?,* Routeledge: London

Beven, K.J. (2012) *Rainfall-Runoff Modelling: The Primer, Second Edition*, John Wiley & Sons Ltd: New York

Beven, K.J., Binley, A.M. (1992) The future of distributed models: Model calibration and predictive uncertainty, *Hydrological Processes*, 6, 279-298

Beven, K. J., Cloke, H. L. (2012) Comment on ‘‘Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water’’ by Eric F. Wood et al., *Water Resources Research*, 48, W01801

Beven, K.J., Freer, J. (2001a) A dynamic TOPMODEL, *Hydrological Processes*, 15(10), 1993-2011

Beven, K.J., Freer, J. (2001b) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems, *Journal of Hydrology*, 249, 11–29

Beven, K.J., Kirkby, M.J. (1979) A physically based, variable contribution area model of basin hydrology*, Hydrological Science Bulletin*, 24, 43-69

Bierkens, M. (2015) Global hydrology 2015: state, trends and directions, Water Resources Research, 51(7), 4923-4947

Bitew, M.M., Gebremichael, M. (2011) Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands, *Hydrology and Earth Systems Science*, 15, 1147-1155

Bondeau, A. et al. (2007) Modelling the role of agriculture fr the 20th century global terrestrial carbon balance, *Global Change Biology*, 13(3), 679–706

Booker, D. J., Woods, R. A. (2014) Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments, *Journal of Hydrology*, 508, 227-239

Boyle, D.P., Gupta, H.V., Sorooshian, S. (2000) Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods, *Water Resources Research*, 36(12), 3663-3674

Breuer, L. et al. (2009) Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use, *Advances in Water Resources*, 32, 129–146

Brown, M.E., Escobar, V., Moran, S., Entekhabi, D., O'Neil, P.E., Njoku, E.G., Doorn, B., Entin, J.K. (2013) NASA's Soil Moisture Active Passive (SMAP) mission and opportunities for applications users, *Journal of Hydrometeorology*, 1125-1127

Burlando, P., Rosso, R. (2002) Effects of transient climate change on basin hydrology. 2. Impacts on runoff variability in the Arno River, central Italy, *Hydrological Processes*, 16, 1177-1199

Burnash, R.J.C., R.L. Ferral, and R.A. McGuire (1973) ' A Generalized Streamflow Simulation System - Conceptual Modeling for Digital Computers', U.S. Department of Commerce, National Weather Service and State of California, Dept. of Water Resources, March

Butts, M.B., Payne, J.T., Kristensen, M., Madsen, H. (2004) An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation, *Journal of Hydrology*, 298(1-4), 242-266

Buytaert, W., Beven, K. (2011) Models as multiple working hypotheses: Hydrological simulation of tropical Alpine wetlands, *Hydrological Processes*, 25, 1784-1799

Carpenter, T.M., Georgakakos, K.P. (2006) Discretisation scale dependencies of the ensemble flow range versus catchment area relationship in distributed hydrologic modelling, *Journal of Hydrology*, 328(1-2), 242-257

Chang, Y., Wu, J, Jiang, G., Kang, Z. (2017) Identification of the dominant hydrological processes and appropriate model structure of a karst catchment through stepwise simplification of a complex conceptual model, *Journal of Hydrology*, 548, 75-87

Chao, Z., Pu, F., Yin, Y., Han. B., Chen. X. (2018) Research on real-time local rainfall prediction based on MEMS sensors, *Journal of Sensors*, Article ID 6184713

Chen, Y., Han. D. (2016) Big data and hydroinformatics, *Journal of Hydroinformatics*, 18(4), 599-614

Chun, K., Wheather, H.S., Onof, C. (2009) Streamflow estimation for six UK catchments under future climate scenarios, *Hydrological Research*, 40(2-3), 96-112

Clark, M.P. et al. (2015a) Improving the representation of hydrologic processes in earth system models, Water Resources Research, 51(8), 5929-5956

Clack, M.P. et al. (2015b) A unified approach for process-based hydrological modelling: 2. Implementation and case studies, *Water Resources Research*, 51(4), 2515-2542

Clark, M.P., Bierkens, M.F.P., Samaniego, L., Woods, R.A., Uijlenhoet, R., Bennett, K.E., Pauwels, V.R.N., Cai, X., Wood, A.W., Peters-Lidard, C.D. (2017) The evolution of process-based hydrological models: historical challenges and the collective quest for physical realism, *Hydrology and Earth System Sciences*, 21, 3427-3440

Clark, M.P., Kavetski, D., Fenicia, F. (2011) Pursuing the method of multiple working hypotheses for hydrological modelling, *Water Resources Research*, 47, W09301

Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T., Hay, L.E. (2008) Framework for understanding structural errors (FUSE): a modular framework to diagnose differences between hydrological models, *Water Resources Research*, 44, W00B02

Clarke, R.T. (2008) Issues of experimental design for comparing the performance of hydrologic models, *Water Resources Research*, 44(1), W01409

Cleugh, H., Leuning, R., Mu, Q., Running, S. (2007) Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sensing of the Environment*, 106(3), 285-304

Cloke, H.L., Hannah, D.M. (2011) Large-scale hydrology: advances in understanding processes, dynamics and models from beyond river basin to global scale, Hydrological Processes, 25(7), 991-995

Coron, L., Andreassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., Hendrickx, F. (2012) Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments, *Water Resources Research*, 48(5), W05552

Coxon, G., Freer, J., Lane, R., Dunne, T., Howden, N. J. K., Quinn, N., Wagener, T., Woods, R. (2018) DECIPHeR v1: Dynamic fluxEs and ConnectIvity for Predictions of HydRology, *Geoscience Model Development*

Coxon, G., Freer, J., Wagener, T., Odoni, N.A., Clark, M. (2014) Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments, *Hydrological Processes*, 28, 6135-6150

Coxon, G., Freer, J., Westerberg, I.K., Wagener, T., Woods, R., Smith, P.J. (2015) A noel framework for discharge uncertainty quantification applied to 500 UK gauging stations, *Water Resources Research*, 51,5531-5546

Cuo, L., Lettenmaier, D.P., Mattheussen, B.V., Storck, P., Wiley, M. (2008) Hydrologic prediction for urban watersheds with the Distributed Hydrology Soil-Vegetation Model, Hydrological Processes, 22(21), 4205-4213

d'Orgeval, T., Polcher, J., de Rosnay, P. (2008) Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes, *Hydrology and Earth Systems Science*, 12, 1387-1401

Dadson, S.J., Ashpole, I., Harris, P., Davies, H.N., Clark, D.B., Blyth, E., Taylor, C.M. (2010) Wetland inundation dynamics in a model of land surface climate: evaluation in the Niger inland delta region, *Journal of Geophysical Research: Atmospheres*, 115, D23114

Dankers, R., Feyen, L. (2008) Climate change impact on flood hazard in Europe: an assessment based on high-resolution climate simulations, *Journal of Geophysical Research,* 113, D19105

Dankers, R., Feyen, L. (2009) Flood hazard in Europe in an ensemble of regional climate scenarios. *Journal of Geophysical Research – Atmospheres*, 114, D16

Das, S., Ashrit, R., Iyengar, G.R., Mohandas, S., Gupta, M.D., George, J.P., Rajagopal, E.N., Dutta, S.K. (2008) Skills of different mesoscale models over Indian region during monsoon season: Forecast errors, *Journal of Earth Systems Science*, 117(5), 603-620

Dawson, C.W., Abrahart, R.J., See, L.M. (2007) HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts, *Environmental Modelling and Software*, 22, 1034-1052

Decharme, B., Alkama, R., Douville, H., Becker, M., Cazenave, A. (2010) Global evaluation of the USBA-TRIP Continental Hydrological System. Part II: Uncertainties in river routing simulation related to flow velocity and groundwater storage, *Journal of Hydrometeorology*, 11(3), 601-617

Decharme, B., Douville, H. (2006) Introduction of a sub-grid hydrology in the ISBA land surface model, *Climate Dynamics*, 26, 65–78.

Decharme, B., Douville, H. (2007) Global validation of the ISBA Sub-Grid Hydrology, *Climate Dynamics*, 29, 21-37

Decharme, B., Martin, E., Faroux, S. (2013) Reconciling soil thermal and hydrological lower boundary conditions in land surface models*, Journal of Geophysical Research: Atmospheres*, 118, 7819-7834

Dembele, M., Zwart, S.J. (2016) Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa, *International Journal of Remote Sensing*, 37(17), 3995-4014

Devi, G.K., Ganasri, B.P., Dwarakish, G.S. (2015) A review of hydrological models, *Aquatic Procedia,* 4, 1001-1007

Dezetter, A., Girard, S., Paturel, J., Servat, E. (2008) Simulation of Runoff in West Africa: is there a single data-model combination that produces the best simulation results?, *Journal of Hydrology*, 345(1-4), 203-212

Di Baldassarre, G., Martinez, F., Kalantari, Z., Viglione, A. (2017) Drought and flood in the Anthropocene: feedback mechanisms in reservoir operation, *Earth System Dynamics*, 8, 225-233

Di Baldassarre, G., Montanari, A. (2009) Uncertainty in river discharge observations: a quantitative analysis, *Hydrology and Earth System Sciences*, 13, 913-921

Diallo, I., Giorgi, F., Deme, A., Tall, M., Mariotti, L., & Gaye, A. T. (2016) Projected changes of summer monsoon extremes and hydroclimatic regimes over West Africa for the twenty-first century, *Climate Dynamics*, 47, 3931–3954

Doll, P. (2009) Vulnerability of the impact of climate change on renewable groundwater resources: a global-scale assessment, Environmental Research Letter, 4(3), 035006

Döll, P., Berkhoff, K., Bormann, H., Fohrer, N., Gerten, D., Hagemann, S., Krol, M. (2008) Advances and visions in large-scale hydrological modelling: findings from the 11th Workshop on Large-Scale Hydrological Modelling, *Advances in Geoscience*, 18, 51–61

Döll, P., Douville, H., Güntner, A., Müller Schmied, H., Wada, Y. (2016) Modeling freshwater resources at the global scale: Challenges and prospects, *Survey of Geophysics*, 37, 195–221

Döll, P., Fiedler, K. (2008) Global-scale modeling of groundwater recharge, *Hydrology & Earth Systems Science*, 12, 863–885

Doll P, Fiedler K, Zhang J (2009) Global-scale analysis of river flow alterations due to water withdrawals and reservoirs*, Hydrology and Earth Systems Science*, 13, 2413–2432

Doll, P., Kaspar, F., Lehner, B. (2003) A global hydrological model for deriving water availability indicators: model tuning and validation, *Journal of Hydrology*, 270, 105-134

Döll, P., Siebert, S. (2002) Global modeling of irrigation water requirements, *Water Resources Research,* 38(4), 801-810

Domeneghetti, A., Castellarin, A., Brath, A. (2012) Assessing rating-curve uncertainty and its effect on hydraulic model calibration, *Hydrology and Earth Systems Sciences*, 16, 1191-1202

Dottori, F., Martina, M.L.V., Todini, E. (2009) A dynamic rating curve approach to indirect discharge measurement, *Hydrology and Earth System Sciences*, 13, 847-863

Duan, Q., et al. (2006) The Model Parameter Estimation Experiment (MOPEX): An overview of science and strategy and major results from the second and third workshops, *Journal of Hydrology*, 320, 3-17

Dugo, M. P., Gao, F. (2011) Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery, *Hydrology and Earth System Sciences*, 15(1), 223

Dunn, S.M., Freer, J., Weiler, M., Kirkby, M.J., Seibert, J., Quinn, P.F., Lischeid, G., Tetzlaff, D., Soulsby, C. (2008) Conceptualisation in catchment modelling: Simply learning?, *Hydrological Processes*, 22, 2389-2393

Entekhabi, D. et al. (2010) The Soil Moisture Active Passive (SMAP) Mission, *Proceedings of the IEEE*, 98(5), 704-716

Euser, T., Winsemius, H.C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., Savenije, H. (2013) A framework to assess the realism of model structures using hydrological signatures, *Hydrology and Earth System Sciences*, 17, 1893-1912

Farmer, D., Sivapalan, M., Jothityangkoon, C. (2003), Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis, *Water Resources Research*, 39(2), 1035

Farr, T. G., et al. (2007), The Shuttle Radar Topography Mission, *Review of Geophysics*, 45, RG2004

Fekete, B., Vorosmarty, C.J., Roads, J., Willmott, C. (2004) Uncertainties in precipitation and their impacts on runoff estimates, *Journal of Climate*, 17, 294-304

Fenicia, F., Kavetski, D., Savenije, H.H.G. (2011) Elements of a flexible approach for conceptual hydrological modelling: 1. Motivation and theoretical development, *Water Resources Research*, 47, W11510

Fenicia, F., Kavetski, D., Savenije, H.H.G., Clark, M.P., Schoups, G., Pfister, L., Freer, J. (2014) Catchment properties, function, and conceptual model representation: Is there correspondence?, *Hydrological Processes*, 28, 2451-2467

Fenton, J. D., Keller, R. J. (2001) The calculation of streamflow from measurements of stage, Technical Report 01/6, Cooperative Research Centre for Catchment Hydrology, Melbourne

Fisher, J.B., Tu, K.P., Baldocchi, D.D. (2008) Global estiamtes of the land-atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites, *Remote Sensing of Environment*, 112, 901-919

Fontes, J.-C., J. N. Andrews, W. M. Edmunds, A. Guerre, and Y. Travi(1991), Paleorecharge by the Niger River (Mali) deduced from ground-water geochemistry, *Water Resources Research*, 27, 199–214

Freer, J., Beven, K.J. and Peters, N.E. (2003) Multivariate seasonal period model rejection within the generalised likelihood uncertainty estimation procedure. In: Calibration of watershed models. Duan, Q., Gupta, H., Sorooshian, S., Rousseau, A.N. and Turcotte, R. (eds), pp. 69-88, AGU, Water Science and Application Series, Washington.

Freer, J., McMillian, H., McDonnell, J.J, Beven, K.J. (2004) Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, *Journal of Hydrology*, 291(3-4), 254-277

Funk, C. et al (2015) The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes, *Nature – Scientific Data*, 2, 150066

Garamboi, P.A., Roux, H., Larnier, K., Labat, D., Dartus, D. (2015) Characterisation of catchment behaviour and rainfall selection for flash flood hydrological model calibration: catchment of the eastern Pyrenees, *Hydrological Sciences Journal*, 60(3), 424-447

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., Sitch, S. (2004) Terrestrial vegetation and water balance – hydrological evaluation of a dynamic global vegetation model, *Journal of Hydrology*, 286(1-4), 249-270

Gerten, D., Schaphoff, S., Lucht, W. (2007) Potential future changes in water limitation of the terrestrial biosphere, *Climatic Change*, 80, 277–299

Getirana, A.C.V., Bonnet, M.P., Rotunno, Filho., O.C., Collischonn, W., Guyot, J.L., Seyler, F., Mansur, W.J. (2010) Hydrological modelling and water balance of the Negro River basin: evaluation based on *in situ* and spatial altimetry data, *Hydrological Processes*, 24, 3219–3236

George Golitzen, K; Andersen, I; Dione, O, Jarosewich-Holder, M. (2005) *The Niger River Basin: A Vision for Sustainable Management. Directions in Development*. Washington, DC: World Bank

Gong, W., Gupta, H.V., Yang, D., Sricharan, K., Hero, A.O. (2013) Estimating epistemic and aleatory uncertaitnies during hydrologic modelling: an information theoretic approach, *Water Resources Research*, 49, 2253-2273

Gosling, S.N., Arnell, N.W. (2011) Simulating current global river runoff with a global hydrological model: model revisions, validation and sensitivity analysis, *Hydrological Processes*, 25(7), 1129-1145

Gosling, S.N., Bretherton, D., Haines, K., Arnell, N.W. (2010) Global hydrology modelling and uncertainty: running multiple ensembles with a campus grid, *Philosophical Transactions of the Royal Society*, 368, 4005-4021

Gourley, J.J., Vieux, B.E. (2006) A method for identifying sources of model uncertainty in rainfall-runoff simulations, *Journal of Hydrology*, 327(1-2), 68-80

Granata, F., Gargano, R., de Marinis, G. (2016) Support vector regression for rainfall-runoff modelling in urban drainage: a comparison with EPA's storm water management mdoel, *Water*, 8(3), 69

Gudmundsson, L., Wagener, T., Tallaksen, I.M., Engeland, K. (2012) Evaluation of nice large-scale hydrological models with respect to the seasonal runoff climatology in Europe, Water Resources Research, 48, W11504

Guntner, A. (2008) Improvement of global hydrological models using GRACE data, *Surveys in Geophysics*, 29(4-5), 375-397

Guo, J., Liang, X., Leung, L. (2004) Impacts of different precipitation data sources on water budgets, *Journal of Hydrology*, 298(1-4), 311-334

Gupta, H.V., Kling, H. (2011) On typical range, sensitivity and normalisation of mean squared error and Nah-Sutcliffe Efficiency type metrics, *Water Resources Research*, 47, W10601

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F. (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, *Journal of Hydrology*, 377, 80-91

Gupta, H.V., Perrin, C. Bloschl, G., Montanari, A., Kumar, R., Clark, M., Andreassian, V. (2014) Large sample hydrology: A need to balance depth and breadth, *Hydrology and Earth System Science*, 18, 463-477

Gupta, H. V., Sorooshian, S. Yapo, P.O. (1999) Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration, *Journal Hydrologic Engineering*, 4(2), 135-143

Gupta, H.V., Wagener, T., Liu, Y. (2008) Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrological Processes*, 22, 3802-3813

Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., Vitvar, T. (2003) A comparative study in modelling runoff and its components in tow mountainous catchments, *Hydrological Processes*, 17(2), 297-311

Haddeland, I. et al. (2011) Multimodel estimate of the global terrestrial water balance: Setup and first results, *Journal of Hydrometeorology*, 12(5), 869-884

Hagemann, S., Göttel, H., Jacob, D., Lorenz, P., Roeckner, E. (2008) Improved regional scale processes reflected in projected hydrological changes over large European catchments, *Climate Dynamics*, 32, 767–781

Hamman, J.J., Nijssen, B., Bohn, T.J., Gergel, D.R., Mao, Y. (2018) The Variable Infiltration Capacity model version 5 (VIC-5): infrastructure improvements for new applications and reproducibility, *Geoscientific Model Development*, 11, 3481-3496

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K. (2008a) An integrated model for the assessment of global water resources – Part 1: Model description and input meteorological forcing, *Hydrology and Earth System Sciences*, 12, 1007–1025

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K. (2008b) An integrated model for the assessment of global water resources – Part 2: Applications and assessments, Hydrology and Earth System Sciences, 12, 1027–1037

Hannah, D.M., Demuth, S., van Lanen, H.A.J., Looser, U., Prudhomme, C., Rees, G., Stahl., K., Tallaksen, L.M. (2011) Large-scale river flow archives: importance, current status and future needs, *Hydrological Processes*, 25(7), 1191-1200

Hargreaves, G.H., Samani, Z.A. (1982) Estimating potential evapotranspiration, *Journal of the Irrigation and Drainage Division*, 108 (3), 225–230

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H. (2014) Updated high-reolution grids of monthly climatic observations – the CRU TS3.10 dataset, International Journal of Climatology, 34(3), 623-642

He, X., Y. Wada, N. Wanders, and J. Sheffield, 2017: Human water management intensifies hydrological drought in California, *Geophysical Research Letter*, 44(4), 1777-1785

Hingray, B., Schaefli, B., Mexghani, A., Hamdi, Y. (2010) Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments, *Hydrological Sciences Journal*, 55(6), 1002-1016

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., Kanae, S. (2013) Global flood risk under climate change, *Nature Climate Change*, 3, 816–821

Hoekstra, A.Y., Chapagain, A.K. (2007) Water footprints of nations: Water use by people as a function of their consumption pattern, *Water Resources Management*, 21, 35–48

Hollander, H.M. et al. (2009) Comparative predictions of discharge from an artificial catchment (Chicken Creek) using sparse data, *Hydrology and Earth System Sciences*, 13, 2069-2094

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T. (2014) The global precipitation measurement mission, Bulletin of the American Meteorological Society, 95(5), 701-722

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H.H.G, Gascuel-Odoux, C. (2014) Process consistency in models: The importance of system signatures, expert knowledge, and process complexity, *Water Resources Research*, 50(9), 7445-7469

Huang, Y., Demuth, S. (2010) *FRIEND—A Global Perspective 2006-2009*, UNESCO: Koblenz

Huang, S. et al. (2017) Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide, *Climate Change*, 141(3), 381-397

Hublart, P., Ruelland, D., Dezetter, A., Jourde, H. (2015) Reducing structural uncertainty in conceptual hydrological modelling in the semi-arid Andes, Hydrology and Earth System Sciences, 19, 2295-2314

Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, D.B. Wolff, (2007) The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale, *Journal of Hydrometeorology*, 8(1), 38-55.

Huffman, G.J., R.F. Adler, D.T. Bolvin, E.J. Nelkin, (2010) The TRMM Multi-satellite Precipitation Analysis (TMPA). Chapter 1 in Satellite Rainfall Applications for Surface Hydrology, F. Hossain and M. Gebremichael, Eds. Springer Verlag, ISBN: 978-90-481-2914-0, 3-22.

Hundecha, Y., Arheimer, B., Donnelly, C., Pechlivanidis, I. (2016) A regional parameter estimation scheme for a pan-European multi-basin model, *Journal of Hydrology: Regional Studies*, 6, 90-111

Hunger, M., Doll, P. (2008) Value of river discharge data for global-scale hydrological modelling, *Hydrology and Earth System Sciences*, 12, 841-861

Hunter, N.M., Bates, P.D., Horritt, M.S., De Roo, A.P.J., Werner, M.G.F. (2005) Utility of different data types for calibrating flood inundation models within a GLUE framework, *Hydrology and Earth System Sciences*, 9, 412-430

Hunter, N.M., Bates, P.D., Horritt, M.S., Wilson, M.D. (2007) Simple spatially-distributed models for predicting flood inundation: a review, *Geomorphology*, 90(3-4), 208-225

Jain, S., Sudheer, K.P. (2008) Fitting of hydrologic models: a close look at the Nash-Sutcliffe index, *Journal of Hydrologic Engineering*, 13(10), 981-986

Jalbert, J., Mathevet, T., Favre, A. (2011) Temporal uncertainty estimation of discharges from rating curves using a variographic analysis, *Journal of Hydrology*, 397(1-2), 83-92

Jimenez, C. et al., (2011) Global intercomparison of 12 land surface heat flux estimates, *Journal of Geophysical Research*, 116, D02102

John, D.M., Lévêque, C., Newton, L.E. (1993) Western Africa. In: D.F. Whigham, D. Dykyjová, and S. Hejný, eds. *Wetlands of the world: inventory, ecology and management*. Vol. 1. Dordrecht: Kluwer, 47–78

Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J., Mechler, R., Botzen, W., Bouwer, L., Pflug, G., Rojas, R., Ward, P. (2014) Increasing stress on disaster-risk finance due to large floods, *Nature Climate Change*, 4, 264-268

Kalma, J. D., McVicar, T. R., McCabe, M. F. (2008). Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data, *Surveys in Geophysics*, 29(4–5), 421–469

Kampf, S., Burges, S.J. (2007) A framework for classifying and comparing distributed hillslope and catchment hydrologic models, *Water Resources Research*, 43(5), W05423

Kauffeldt, A., Wetterhall, F., Pappenberger, P., Salamon, P., Thielen, J. (2016) Technical rebiew of large-scale hydrological models for implementation in operational flood forecasting schemes on continental levels, *Environmental Modelling & Software,* 75, 68-76

Kavetski, D., Clark, M.P. (2010) Ancient numerical daemons of conceptual hydrological modelling: 2. Impact of time stepping schemes on model analysis and prediction, *Water Resources Research*, 46(10), W10511

Kavetski, D., F. Fenicia, and M. P. Clark (2011), Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment, *Water Resources Research*, 47, W05501

Khakbaz, B., Imam, B., Hsu, K., Sorooshian, S. (2009) From lumped to distributed via semi-distributed: calibration strategies for semi-distributed hydrologic models, *Journal of Hydrology*, 418-419, 61-77

Kirchner, J.W. (2006) Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resources Research*, 42, W03S04

Kidd, C., Bauer, P., Turk, J., Huffman, G. J., Joyce, R., Hsu, K.-L., Braithwaite, D. (2011) Intercomparison of High-Resolution Precipitation Products over Northwest Europe, Journal of Hydrometeorology, 13, 67–83

Kidd, C., Huffman, G. (2011) Global precipitation measurement, *Meteorological Applications*, 18(3), 334-353

Kobierska, F., Jonas, T., Zappa, M., Bavay, M., Magnusson, J., Bernasconi, S.M. (2013) Future runoff from a partly glacierised watershed in Central Switzerland: a 2 model approach, *Advances in Water Resources*, 55, 204-214

Kokkonen, T., Kouvusalo, H., Karvonen, T. (2001) A semi-distributed approach to rainfall-runoff modelling – a case study in a snow affected catchment, *Environmental Modelling and Software*, 16(5), 481-493

Konikow, L. (2011) Contribution of global groundwater depletion since 1900 to sea-level rise, Geophysical Research Letters, 38(17), *Hydrology and Land Surface Studies*, L17401

Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D. (2004) Hydrology laboratory research modelling system (HL-RMS) of the US National weather service, *Journal of Hydrology*, 291, 297-318

Krueger, T., Freer, J., Quinton, J.N., Macleod, C.J.A., Bilotta, G.S., Brazier, R.E., Butler, P., Haygarth, P.M. (2010) Ensemble evaluation of hydrological model hypotheses*, Water Resources Research*, 46, W07516

Krause, P., Boyle, D.P., Base, F. (2005) Comparison of different efficiency criteria for hydrological model assessment, *Advances in Geosciences*, 5, 89-97

Kurc, S., Small, E.E. (2004) Dynamics of evapotranspiration in semiarid grassland and shrubland ecosystems during the summer monsoon season, central New Mexico, *Water Resources Research*, 40, W09305

Le Moine, N., Andréassian, V., Perrin, C., Michel, C. (2007), How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, *Water Resources Research*, 43, W06428

Lee, H., McIntyre, N., Wheater, H., Young, A., (2005) Selection of conceptual models for regionalisation of the rainfall-runoff relationship, *Journal of Hydrology*, 312, 125-147

Lee, T., Srinivasan, R., Moon, J., Omani, N. (2011) Estimation of fresh water inflow to bays from gauged and ungauged watersheds, *Applied Engineering in Agriculture*, 27(6), 917-923

Lehner, B., et al. (2011) High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9(9), 494–502

Lettenmaier, D.P., Famiglietti, J.S. (2006) Hydrology: water from on high, Nature, 444(7119), 562-563

Ley, R., Hellebrand, H., Casper, M.C., Fenicia, F. (2016) Is catchment classification possible by means of multiple model structures? A case study based on 99 catchments in Germany, *Hydrology*, 3 (22)

Li, L., Diallo, I., Xu, C., Stordal, F. (2015) Hydrological projections under climate change in the near future by RegCM4 in Southern Africa using a large-scale hydrological model, *Journal of Hydrology*, 528, 1-16

Li, S., Kang, S., Zhang, L., Zhang, J., Du,. T., Tong, L., Ding, R. (2016) Evaluation of six potential evapotranspiration models for estimating crop potential and actual evapotranspiration in arid regions, *Journal of Hydrology*, 543, 450-461

Liang, X., D. P., Lettenmaier, E. F., Wood, S. J. Burges, (1994) A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs, *Journal of Geophysical Research.*, **99**(D7), 14415-14428

Liang, X., Wood, E. F., Lettenmaier, D. P. (1996) Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications, *Global and Planetary Change*, 13, 195-206

Liebscher, H. (1993) Hydrology for the water management of large river basins, *Hydrological Sciences Journal*, 38(1), 1-13

Liou, Y. A., Kar, S. K. (2014) Evapotranspiration estimation with remote sensing and various surface energy balance algorithms – A review, *Energies*, 7(12), 2821–2849

Liersch, S. et al. (2013) Vulnerability of rice production in the Inner Niger Delta to water resources management under climate variability and change, *Environmental Science & Policy*, 34, 18-33

Liu, Q., Yang, Z., Cui, B. (2009) Spatial and temporal variability of annual precipitation during 1961–2006 in Yellow River Basin, China, *Journal of Hydrology*, 361(3–4), 330–338

Liu, X., Tang, Q., Mu, M., Gerten, D., Gosling, S., Masaki, Y., Satoh, Y., Wada, Y. (2017) Multimodel uncertainty changes in simulated river flows induced by human impact parameterizations, *Environmental Research Letters*, 12(2), 025009

Lobligeois, F., Andreassian, V., Perrin, C., Tabary, P., Loumagne, C. (2014) When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, *Hydrology and Earth System Sciences*, 18, 575-594

López López, P., Wanders, N., Schellekens, J., Renzullo, L.J., Sutanudjaja, E.H., Bierkens, M.F.P. (2016) Improved large-scale hydrological modelling through the assimilation of streamflow and downscales satellite soil moisture observations, *Hydrology and Earth System Sciences*, 20(7), 3059–3076

Louvet, S., et al. (2011) Variabilité spatio-temporelle passée et future de la pluie sur le basin du Bani en Afrique de l’Ouest. In: S. Franks, E. Boegh, E. Blyth, D.M. Hannah, and K.K. Yilmaz, eds. Hydro-climatology: variability and change, (Proceedings of symposium J-H02 held during IUGG2011 in Melbourne, Australia, July 2011). Wallingford: IAHS Publication 344, 125–130

Mahe, G., Bamba, F., Soumaguel, A., Orange, D., Olivry, J.C. (2009) Water losses in the inner delta of the River Niger: water balance and flooded area, *Hydrological Processes,* 23(22), 3157-3160

Mahe, G., Lienou, G., Descroix, L., Khomsi, K. (2013) The river of Africa: Witness of climate change and human impact on the environment, *Hydrological Processes*, 27(15), 2105-2114

Martens, B., et al. (2017) GLEAM v3: satellite-baed land evaporation and root-zone soil moisture, *Geoscientific Model Development*, 10(5), 1903–1925

Maxwell, R.M., Miller, N.L. (2005) Development of a coupled land surface and groundwater model, *Journal of Hydrometeorology*, 6, 233-247

McCuen, R. H., Knight, Z., Cutter, A.G. (2006) Evaluation of the Nash-Sutcliffe Efficiency Index, *Journal of Hydrologic Engineering*, 11(6), 597-602

McDonnell, J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. Hooper, J. Kirchner, and M. Roderick, (2007) Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resources Research*, 43

McDonnell, J., Woods, R. (2004) On the need for catchment classification, *Journal of Hydrology*, 299, 2-3

McEnery, J., Ingram, J., Duan, Q., Adams, T., Anderson, L. (2005) NOAA'S Advanced hydrologic prediction service: building pathways for better science in water forecasting, BAMS, 86(3), 375-386

McMahon, T.A., Peel, M.C., Lowe, L., Srikanthan, R, McVicar, T.R. (2013) Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, *Hydrology and Earth System Sciences*, 17, 1331-1363

McMillan, H.K., Booker, D.J., Cattoen, C. (2016) Validation of a national hydrological model, *Journal of Hydrology*, 541 (Part B), 800-815

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., Clark, M. (2010) Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions, *Hydrological Processes*, 24, 1270-1284

McMillan H, Jackson, B., Clark, M., Kavetski, D., Wood, R. (2011) Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models, *Journal of Hydrology*, 400(1-2), 83-94

McMillan, H.K., Westerberg, I.K. (2015) Rating curve estimation under epistemic uncertainty, *Hydrological Processes*, 29, 1873-1882

McMillan, H., Westerberg, I., Branger, F. (2017) Five guidelines for selecting hydrological signatures, *Hydrological Processes*, 31(26), 4757-4761

McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N. M., Mescherskaya, A. V., Kruger, A. C., Rehman, S., Dinpashoh, Y. (2012) Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation, *Journal of Hydrology*, 416, 182–205

Menne, M., Durre, I., Vose, R., Gleason, B.E., Houston, T.G. (2012) An overvuiew of the Global Historical Climatology Network-Daily Database*, Journal of Atmospheric and Oceanic Technology*, 29(7), 897-910

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., Cox, P. M. (2009) Impact of changes in diffuse radiation on the global land carbon sink, *Nature*, 458(7241), 1014–1017

Merz, R., Bloschl, G. (2004) Regionalisation of catchment model parameters, *Journal of Hydrology*, 287, 95-123

Miralles, D.G., et al. (2016) The WACMOS-ET project – Part 2: Evaluation of global terrestrial evaporation data sets, *Hydrology and Earth System Sciences*, 20(2), 823-842

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., Dolman, A. J. (2011) Global land-surface evaporation estimated from satellite-based observations, *Hydrology Earth System Sciences*, 15, 453–469

Mitchell, T.D., Jones, P.D. (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *International Journal of Climatology*, 25, 693–712

Mockler, E.M., Chun, K.P., Sapriza-Azuri, M., Bruen, M., Wheather, H.S. (2016) Assessing the relative importance of parameter and forcing uncertainty and their interactions in conceptual hydrological model simulations*, Advances in Water Resources*, 97, 299-313

Moges, E., Demissie, Y., Li, H. (2016) Hierarchical mixture of experts and diagnostic modelling approach to reduce hydrologic model structural uncertainty, Water Resources Research, 52(4), 2551-2570

Molini, A., La Barbera, P., Lanza, L.G., Stagi, L. (2001) Rainfall intermittency and the sampling error of tipping-bucket rain gauges, *Physics and Chemistry of the Earth*, 26(10–12), 737–742

Montanari, A. et al. (2013) “Panta Rhei – Everything Flows”: Change in hydrology and society – the IAHS Scientific Decade 2013-2022, *Hydrological Sciences Journal*, 58(6), 1256-1275

Monteith, J.L. (1965) Evaporation and the Environment, *19th Symposia of the Society for Experimental Biology*, 19, 205-234

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, *Transactions of the ASABE*, 50(3), 885-900

Moulin, L., Gaume, E., Obled, C.(2009)Uncertainties on mean areal precipitation: assessment and impact on streamflow simulations, *Hydrology and Earth Systems Science*,13, 99–114.

Moussa, R., Chahinian, N. (2009) Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events, *Hydrology and Earth System Sciences*, 13, 519-535

Moyeed, R.A., Clarke, R.T. (2005) The use of Bayesian methods for fitting rating curves, with case studies, *Advances in Water Resources*, 28, 807-818

Mu, Q., et al. (2007) Development of a global evapotranspiration algorithm based on MODIS and global meteorology data, *Remote Sensing of Environment*, 111 (4), 519–536

Mu, Q., Zhao, M., Running, S. (2011) Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sensing of the Environment*, 115(8), 1781-1800

Mueller, B., et al. (2011) Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophysical Research Letters*, 38, L06402

Nair, A. S., Indu, J. (2017) Performance assessment of Multi-Source Weighted-Ensemble Precipitation (MSWEP) product over India, *Climate*, 5, 2

Nash, J.E., Sutcliffe, J.V. (1970) River flow forecasting through conceptual models part 1 –A discussion of principles, *Journal of Hydrology*, 10(3), 282-290

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. (2009) Soil and Water Assessment Tool. Theoretical Documentation Version 2005. Grassland, Soil And Water Research Laboratory Agricultural Research Service, Texas 76502, Blackland Research Center, Texas Agricultural Experiment Station, Texas 76502 USA

Newman, A.J., Clark, M.P., Craig, J. (2015) Gridded ensemble precipitation and temperature estimates for the contiguous United States, *Journal of Hydrometeorology*, 16(6), 2481-2500

Nkamdjou, L.S., Bedimo, J.B. (2008) Critique des Donnees Hydrometriques du Bassin du Niger: Rapport Preliminaire; Niger Basin Authority: Niamey, Niger, p. 28

Oki, T., Kanae, S. (2006) Global hydrological cycles and world water resources, *Science*, 313(5790), 1068-1072

Olomoda, Ibraheem A. (2002): Impact of climatic change on river Niger. Niger Basin Authority. http://www.riversymposium.com/2006/index.php?element=06OLOMODAIbraheem

Orlowsky, B., Seneviratne, S.I. (2013) Elusive drought: uncertainty in observed trends and short- and long-term CMIP5 projections, *Hydrology and Earth System Sciences*, 17, 1765-1781

Orth, R., Staudinger, M., Seneviratne, S.I., Seibert, J., Zappa, M. (2015) Does model performance improve with complexity? A case study with three hydrological models, *Journal of Hydrology*, 523, 147-159

Oudin, L., Andreassian, V., Perrin, C., Michel, C., Moine, N.L. (2008) Spatial proximity, physical similarity, regression and ungauged catchments: a comparison of regionalisation approaches based on 913 French catchments, *Water Resources Research*, 44(3), W03413

Oudin L, et al. (2005) Which potential evapotranspiration input for a lumped rainfall-runoff model? (Part 2): Towards a simple and efficient potential evapotranspiration sensitivity analysis and identification of the best evapotranspiration and runoff options for rainfall-runoff modelling, *Journal of Hydrology*, 303(1-4), 290-206

Oudin, L., Perrin, C., Mathevet, T., Andreassian, V., Michel, C. (2006) Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models, *Journal of Hydrology*, 320(1-2), 62-83

Pappenberger, F., Bartholmes, J., Thielen, J., Cloke, H.L., de Roo, A., Buizza, R. (2008) New dimensions in early flood warning across the globe using GRAND ensembles, *Geophysical Research Letters*, 35, L10404

Pappenberger, F., Beven, K. J., Hunter, N., Gouweleeuw, B., Bates, P., de Roo, A., Thielen, J. (2005) Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS), *Hydrology and Earth System Sciences*, 9, 381–393

Park, J., Botter, G., Jawitz, J.W., Rao, P.S.C. (2014) Stochastic modeling of hydrologic variability of geographically isolated wetlands: effects of hydro-climatic forcing and wetland bathymetry, *Advances in Water Resources*, 69, 38–48

Paturel, J.E., Barrau, C., Mahe, G., Dezetter, A., Servat, E. (2007) Modelling the impact of climatic variability on water resources in West and Central Africa from a non-calibrated hydrological model, *Hydrological Sciences Journal*, 52(1), 38-48

Pechlivanidis, I.G., Jackson, B.M., McIntyre, N.R., Wheater, H.S. (2011) Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications, *Global NEST Journal*, 13(3), 193-214

Pedinotti, V., Boone, A., Decharme, B., Cretaux, J.F., Mognard, N., Panthou, G., Papa, F., Tanimoun, B.A. (2012) Evaluation of the ISBA-TRIP continental hydrological system over the Niger Basin using in situ and satellite derived datasets, *Hydrology and Earth System Sciences*, 16, 1745-1773

Peel, M.C., McMahon, T.A., Finlayson, B.L. (2010) Vegetation impact on mean annual evapotranspiration at a global catchment scale, *Water Resources Research*, 46(9), W09508

Pendergrass, A. G., Hartmann, D. L. (2014) Two modes of change of the distribution of rain, *Journal of Climate*, 27, 8357–8371

Perrin, C., Dilks, C., BArlund, I., Payan, J.L., Andreassian, V. (2006) Use of simple rainfall-runoff models as a baseline for the benchmarking of the hydrological component of complex catchment models, *Large River – Arch. Hydrobiol. Suppl.*, 17(1-2), 75-96

Perrin, C., Michel, C., Anreassian, V. (2001) Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *Journal of Hydrology*, 242, 275-301

Petersen-Øverleir, A. (2006) Modelling stage—discharge relationships affected by hysteresis using the Jones formula and nonlinear regression, *Hydrological Sciences Journal*, 51(3), 365-388

Petersen-Øverleir, A., Soot, A., Reitan, T. (2009) Bayesian rating curve inference as a streamflow data quality assessment tool, *Water Resource Management*, 31, 1835-1842

Pokhrel, Y. N., Koirala, S., Yeh, P. J.-F., Hanasaki, N., Longuevergne, L., Kanae, S., Oki, T. (2015) Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts, *Water Resources Research*, 51, 78–96

Priestley, C.H.B., Taylor, R.J. (1972) On the assessment of the surface heat flux and evaporation using large-scale parameters, *Monthly Weather Review*, 100, 81-92

Prudhomme, C. et al. (2014) Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment, *PNAS*, 111(9), 3262-3267

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., Andreassian, V. (2011) A downward structural sensitivity analysis of hydrological mdoels to improve low-flow simulation, *Journal of Hydrology*, 411(1-2), 66-76

Pushpalatha, R., Perrin, C., Moine, N.L., Andreassian, V. (2012) A review of efficiency criteria suitable for evaluating low-flow simulations, *Journal of Hydrology*, 420-421, 171-182

Qu, Y. (2004) An integrated hydrologic model for multi-process simulation using semidiscrete finite volume approach. Ph.D. thesis, The Pennsylvania State University

Raje, D., Priya, P., and Krishnan, R. (2013) Macroscale hydrological modelling approach for study of large-scale hydrologic impacts under climate change in Indian River basins, Hydrological Processes, 28, 1874–1889

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.J. (2004) Overall distributed model intercomparison project results, *Journal of Hydrology* 298, 27–60

Renard, B., Kavetski, D., Thyer, M., Kuczera, G., Franks, S.W. (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resources Research*, 46

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G., Franks, S.W. (2011) Toward a reliable decomposition of predictive uncertainty in hydrological modelling: characterising rainfall errors using conditional simulation, *Water Resources Research*, 47(11), W11516

Reusser, D.E., Blume, T., Schaefli, B., Zehe, E. (2009) Analysing the temporal dynamics of model performance for hydrological models, *Hydrology and Earth System Sciences*, 13, 999-1018

Ritter, A., Munoz-Carpena, R. (2013) Performance evaluation of hydrological models: statistical significance for reducing subjectivity in goodness-of-fit assessments, *Journal of Hydrology*, 480, 33-45

Rodell, M. et al. (2004) The Global Land Data Assimilation System, *Bulletin of the American Meteorological Society*, 85(3), 381-394

Rost, S., et al. (2008) Agricultural green and blue water consumption and its influence on the global water system, *Water Resources Research*, 44, W09405

Saha, S., et al. (2010) The NCEP climate forecast system reanalysis, *Bulletin of the American Meteorological Society*, 91(8), 1015-1056

Salamon, P. Feyen, L. (2010) Disentangling uncertainties in distributed hydrological modelling using multiplicative error models and sequential data assimilation, *Water Resources Research*, 46(12)

Samain, B., Pauwels, V. (2013) Impact of potential and (scintillometer-based) actual evapotranspiration estimates on the performance of a limped rainfall-runoff model, *Hydrology and Earth System Sciences*, 17, 4525–4540

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P.A., Carrillo, G. (2011) Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, *Hydrology and Earth System Sciences*, 15, 2895-2911

Schaefli, B., Gupta, H.V. (2007) Do Nash values have values?, *Hydrological Processes*, 21, 2075, 2080

Schellekens, J. et al. (2017) A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, *Earth System Science Data*, 9, 389-413

Schuol, J., Abbaspour, K. C. (2006) Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. Advances in Geosciences*, European Geosciences Union*, 9, 137-143

Schoups, G., N. C. van de Giesen, and H. H. G. Savenije (2008), Model complexity control for hydrologic prediction, *Water Resources Research*, 44, W00B03

Shi, T., Guan, D., Wu, J., Wang, A., Jin, C., Han, S. (2008) Comparison of methods for estimating evapotranspiration rate of dry forest canopy: eddy covariance, Bowen ration energy balance, and Penman-Montheith equation, *Journal of Geophysical Research*, 113, D19116

Shrestha, R., Tachikawa, Y., Takara, K. (2006) Input data resolution analysis for distributed hydrological modelling, *Journal of Hydrology*, 319, 36-50

Shukla, J., Mintz, Y. (1982) Influence of land-surface evapotranspiration on the Earth's climate, *Science*, 215(4539), 1498-1501

Seibert, J., Bishop, K., Nyberg, L., Rodhe, A. (2011) Water storage in a till catchment: I. Distributed modeling and relationship to runoff, *Hydrological Processes*, 25, 3937–3949

Siebert, S., Doll, P. (2010) Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, *Journal of Hydrology*, 384(3–4), 198–217

Siebert, S., Hoogeveen, J., and Frenken, K. (2006) Irrigation in Africa, Europe and Latin America. Update of the Digital Global Map of Irrigation Areas to Version 4, Frankfurt Hydrology Paper 05, Institute of Physical Geography, University of Frankfurt, Frankfurt am Main, Germany, 134 pp

Siebert, S., Kummu, M., Porkka, M., Doll, P., Ramankutty, N., Scanlon, B.R. (2015) A global data set of the extent of irrigated land from 1900 to 2005, *Hydrology and Earth System Sciences*, 19, 1521-1545

Simmons, A.J., Willett, K.M., Jones, P.D., Thorne, P.W., Dee, D.P. (2010) Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets, *Journal of Geophysical Research: Atmospheres*, 151(D1), D01110

Singh, V.P. (1995) *Computer models of watershed hydrology*, Water Recourses Publications: LLC, USA

Singh, V.P. (1997) Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph, *Hydrological Processes*, 11(12), 1649-1669

Singh, V.P., Frevert, D.K. (2006) *Watershed Models*, Taylor and Francis Group: USA

Singh, V.P. and Woolhiser, D.A. (2002) Mathematical modelling of watershed hydrology, *Journal of Hydraulic Engineering*, 7, 270-292

Sivapalan, M., Bloschl. G., Merz, R., Gutknecht, D. (2005) Linking flood frequency to long-term water balance: incorporating effects of seasonality, *Water Resources Research*, 41, W06012

Sivapalan, M. et al. (2003) IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences*, Hydrological Sciences Journal*, 48(6), 857-880

Sivapalan, M., Konar, M., Srinivasan, V., Chhatre, A., Wutich, A., Scott, C.A., Wescoat, J.L., Rodriguez-Iturbe, I. (2014) Socio-hydrology: Use-inspired water sustainability science for the Anthropocene, *Earth's Future*, 2, 225-230

Sivapalan, M., Savenije, H.H.G., Bloschl, G. (2012) Socio-hydrology: A new science of people and water, *Hydrological Processes*, 26, 1270-1276

Sivapalan, M. (2015) Debates – Perspectives on socio-hydrology: changing water systems and the 'tyranny of small problems' – Socio-hydrology, *Water Resources Research*, 51(6), 4795-4805

Slater, A. G., et al. (2001), The representation of snow in land-surface schemes: Results from PILPS 2 (d), *Journal of Hydrometeorology*, 2, 7–25

Smith, M. B., Koren, V., Reed, S., Zhang, Z., Zhang, Y., Moreda, F., Cui, Z., Mizukami, N., Anderson, E. A., Cosgrove, B. A. (2012) The distributed model intercomparison project – Phase 2: Motivation and design of the Oklahoma experiments, *Journal of Hydrology*, 418– 419, 3–16

Smith, T.J., Marshall, L.A. (2010) Exploring uncertainty and model predictive performance concepts via a modular snowmelt-runoff modelling framework, *Environmental Modelling and Software*, 25, 691-701

Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S., Zhang, Z., Duan, Q., Moreda, F., Cong, S. (2004) The distributed model intercomparison project (DMIP): Motivation and experiment design, *Journal of Hydrology*, 298, 4–26

Son, K., Sivapalan, M., (2007) Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, *Water Resources Research*, 43, W01415

Sood, A., Smakhtin, V. (2014) Global hydrological models: a review, *Hydrological Sciences Journal*, 60(4), 549-565

Stahl, K., Tallaksen, L.M., Gudmundsson, L., Christensen, J.H. (2011) Streamflow data from small basins: a challenging test to high-resolution regional climate modelling, *Journal of Hydrometeorology*, 11, 1275-1285

Staudinger, M., Stahl, K., Seibert, J., Clark, M.P., Tallaksen, L.M. (2011) Comparison of hydrological model structures based on recession and low flow simulations, *Hydrology and Earth System Sciences*, 15, 3447-3459

Stewart, R.B., Rouse, W.R. (1977) Substantiation of the Priestly and Taylor parameter a = 1.26 for Potential Evaporation in high latitudes, *Journal of Applied Meteorology*, 16(6), 649-650

Strauch, M., Bernhofer, C., Koide, S., Volk, M., Lorz, C., Makeschin, F. (2012) Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation, *Journal of Hydrology*, 414-415, 413-424

Syed, T., Famiglietti, J.S., Rodell, M., Chen, J., Wilson, C.R. (2008) Analaysis of terrestrial water storage changes from GRACE and GLDAS, *Water Resources Research*, 44(2), W02433

Su, Z., Dorigo, W., Fernández-Prieto, D., Van Helvoirt, M. V., Hungershoefer, K., de Jeu, R. D., Schulz, J. (2010). Earth observation water cycle multi-mission observation strategy (WACMOS). *Hydrology and Earth System Sciences Discussions*, 7(5), 7899–7956

Sun, X., Mein, R.G., Keenan, T.D., Elliott, J.F.(2008)Flood estimation using radar and raingauge data, *Journal of Hydrology*, 239, 4–18

Sutcliffe, J.V., Parks, Y.P. (1989) Comparative water balances of selected African wetlands. *Hydrological Sciences Journal*, 34, 49–62

Tadono, T., Takaku, J., Tsutsui, K., Oda, F., Nagai, H. (2015), Status of“ALOS World 3D (AW3D)”global DSM generation, Proceeding 2015 IEEEInternational Geoscience and Remote Sensing Symposium (IGARSS), 3822–3825

Tang, Q., Gao, H., Yeh, P., Oki, T., Su, F. and Lettenmaier, D.P. (2009) Dynamics of terrestrial water storage change from observations and modelling, *Journal of Hydrometeorology*, 11, 156-170

Tapley, B.D., Bettadpur, S., Ries, J.C., Thompson, P.F., Watkins, M.M. (2004) GRACE measurements of mass variability in the earth system, *Science*, 305, 503-505

Tarhule, A., et al. (2015) Exploring temporal hydroclimatic variability in the Niger Basin (1901-2006) using observed and gridded data, *International Journal of Climatology*, 35(4), 520-539

Thompson, J.R., Crawley, A., Kingston, D.G. (2016) GCM-related uncertainty for river flows and inundation under climate change: the Inner Niger Delta, *Hydrological Sciences Journal*, 61, 2325–2347

Thompson, J.R., Crawley, A., Kingston, D.G. (2017) Future river flows and flood extent in the Upper Niger and Inner Niger Delta: GCM-related uncertainty using the CMIP5 ensemble, *Hydrological Sciences Journal*, 62(14), 2239-2265

Tian, Y., Peters-Lidard, C.D. (2010) Real-time bias reduction for satellite-based precipitation estimates, *Journal of Hydrometeorology*, 11, 1275-1285

Tomkins, K.M. (2014) Uncertainty in streamflow rating curves: Methods, controls and consequences, *Hydrological Processes*, 28, 464–481

Trambauer, P., Maskey, S., Werner, M., Pappernberger, F., van Beek, L., Uhlenbrook, S. (2014) Identification and simulation of space-time variability of past hydrological drought events in the Limpopo River basin, southern Africa, *Hydrology and Earth System Sciences*, 18(8), 2925–2942

Troch, P.E. Carrillo, G.A., Heidbuchel, I., Rajagopal, S., Switanek, M., Volkmann, T.H.M., Yaeger, M (2009) Dealing with landscape heterogeneity in watershed hydrology: a review of recent progress toward new hydrological theory, *Geography Compass*, 3(1), 375-392

Troy, T.J., Wood, E.F., Sheffield, J. (2008) An efficient calibration method for continental-scale land surface modelling, *Water Resources Research*, 44(9), W09411

Tuanmu, M.-N. Jetz, W. (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem modelling, *Global Ecology and Biogeography*, 23, 1031–1045

Tyralla, C., Schumann, A.H. (2016) Incorporating structural uncertainty of hydrological models in likelihood functions via an ensemble range approach, *Hydrological Sciences Journal*, 61(9), 1679-1690

Uhlenbrook, S., McDonnell, J., Leibundgut, C. (2003) Runoff generation and implications for river basin modelling, *Hydrological Process*, 17, 197-512

van Beek, L.P.H., Wada, Y., Bierkens, M.F.P. (2011) Global monthly water stress: 1. Water balance and water availability, *Water Resources Research*, 47(7), W07517

van Camp, M., Meurers, B., de Viron, O., Forbriger, T. (2016) Optimized strategy for the calibration of superconducting gravimeters at the one per mille level, *Journal of Geodesy*, 90(1), 91–99

Van Dijk, A. I. J. M. (2010), AWRA Technical Report 3, Landscape Model (Version 0.5) Technical Description, WIRADA/CSIRO Water for a Healthy Country Flagship, Canberra. [Available at http://www.clw.csiro.au/publications/waterforahealthycountry/2010/wfhc-aus-water-resources-assessment-system.pdf.]

Van Dijk, A., Pena-Arancibia, J., Wood, E., Sheffield, J., Beck, H. (2013) Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide, *Water Resources Research*, 49(5), 2729–2746

Van Dijk, A.I.J.M., Renzullo, L.J., Wada, Y., Tregoning, P. (2014) A global water cycle reanalysis (2003-2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble, *Hydrology and Earth System Sciences*, 18, 2955-2973

Van Esse, W.R., Perrin, C., Booij, M.J., Augustijn, D.C.M., Fenicia, F., Kavetski, D., Lobligeois, F. (2013) The influence of conceptual model structure on model performance: a comparative study for 237 French Catchments, *Hydrology and Earth System Sciences*, 17, 4227-4239

Van Loon, A.F., et al. (2016) Drought in the Anthropocene, *Nature Geoscience*, 9, 89-91

Van Loon, A., Van Lanen, H. (2013) Making the distinction between water scarcity and drought using an observation-modelling framework, *Water Resources Research*, 49(3), 1483-1502

Vansteenkiste, T., Tavakoli, M., Ntegeka, V., De Smedt, F., Batelaan, O., Pereira, F., Willems, P. (2014) Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections, *Journal of Hydrology*, 519, 743-755

Vaze, J., Jordan, P., Beecham, R., Frost, A., Summerell, G. (2012) *Guidelines for Rainfall-Runoff Modelling: Towards best practice model application* (pp. 47)

Velaquez, J.A., Anetil, F., Ramos, M.H., Perrin, C. (2011) Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures, *Advances in Geosciences*, 29, 33-42

Veldkamp, T., Eisner, S., Wada, Y., Aerts, J., Ward, P. (2015) Sensitivity of water scarcity events to ENSO-driven climate variability at the global scale, *Hydrology and Earth System Sciences*, 19, 4081-4098

Veldkamp, T., et al. (2018) Human impact parameterisations in global hydrological models improves estimates of monthly discharge and hydrological extremes: a multi-model validation study, *Environmental Research Letters*, 13, 055008

Villarini, G., Smith, J.A., Baeck, M.L., Sturdevant-Rees, P., Krajewski, W.F. (2010) Radar Analyses of Extreme Rainfall and Flooding in Urban Drainage Basins, *Journal of Hydrology*, 381, 266-286

Vogel, R.M., Fennessey, N.M. (1994) Flow-duration curves. I: New interpretation and confidence intervals, *Journal of Water Resources Planning and Management*, 120, 485-504

Voisin, N., Li, H., Ward, D., Huang, M., Wigmosta, M., Leung, L. R. (2013) On an improved sub-regional water resources management representation for integration into earth system models, *Hydrology and Earth System Sciences*, 17, 3605–3622

Vrugt, J.A., Diks, C., Gupta, H., Bouten, W., Verstraten, J. (2005) Improved treatment of uncetaintiy in hydrological modelling: combining the strengths of global optimization and data assimilation, *Water Resource Research*, 41, W01017

Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W., Sorooshian, S. (2003) Effective and efficient algorithm for multiobjective optimisation of hydrologic models, Water Resources Research, 39(8), 1214

Wada, Y., et al. (2016) Modelling global water use for the 21st century: the Water Futures and Solutions (WFaS) initiative and its approaches, *Geoscientific Model Development*, 9, 175-222

Wada, Y., et al. (2017) Human-water interface in hydrological modelling: current status and future direction, *Hydrology and Earth System Sciences*, 21(8), 4169–4193

Wada, Y., Wisser, D., Bierkens, M.F.P. (2014) Global modelling of withdrawal, allocation and consumptive use of surface water and groundwater resources, *Earth System Dynamics*, 5, 15-40

Wagener, T. (2003) Evaluation of catchment models, *Hydrological Processes*, 17(16), 3375–3378

Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V., Sorooshian, S. (2001) A framework for development and application of hydrological models, *Hydrology and Earth System Science*, 5, 13-26

Wagener, T., Gupta, H.V. (2005) Model identification for hydrological forecasting under uncertainty, *Environmental Research and Risk Assessment*, 19(6), 378-387

Wagener, T., McIntyre, N., Lees, M.J., Wheather, H.S., Gupta, H.V. (2003) Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrological Processes*, 17, 455-476

Wagener, T., Lees, M.J., Wheater, H.S. (2002) A toolkit for the development and application of parsimonious hydrological models, in *Mathematical models of large watershed hydrology - Volume 1*, 1 edn, V. P. Singh & D. Frevert, eds., USA, pp. 87-136

Wagener, T., Montanari, A. (2011) Convergence of approaches toward reducing uncertainty in predictions in ungauged basins, *Water Resources Research*, 47(6), W06301

Wagener, T., Wheather, H.S. (2006) Parameter estimation and regionalisation for continuous rainfall-runoff models including uncertainty, *Journal of Hydrology*, 320, 132-154

Walker, W.E., Harremoes, P., Rotmsd, J., van der Sluijs, J.P. van Asselt, M.B.A., Janssen, P., von Krauss, M.P.K. (2003) Defining Uncertainty: A conceptual basis for uncertainty management in model-based decision support, *Integrated Assessment*, 4, 5-17

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P. (2014) The WFDEI meteorological focing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resources Research*, 50(9), 7505-7514

Weedon, G.P. et al. (2011) Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *Journal of Hydrometeorology*, 12, 823-848

Werth, S., White, D., Bliss, D.W. (2017) GRACE detected rise in groundwater in the Sahelian Niger River Basin, *Journal of Geophysical Research: Solid Earth*, 122(10), 459-477

Westerberg, I.K., Guerro, J.L., Younger, P.M., Beven, K.J., Seibert, J., Halldin, S., Freer, J.E. Xu, C.Y. (2011) Calibration of hydrological models using flow-duration curves, *Hydrology and Earth System Sciences*, 15, 2205-2227

Wheater, H.S. (2002) Progress in and prospects for fluvial flood modelling, *Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineering Sciences*, 360, 1409-1431

Wheater, H., Jakeman, A., Beven, K. (1993) Chapter 5 –Progress and directions in rainfall-runoff modelling. In: Jakeman, A.J., Beck, M.B., McAleer, M.J. (eds), Modelling Change in Environmental System, Wiley: Chichester, UK

WHYMAP, & J. Margat (2008).Large aquifer systems of the world, map source: http://www.whymap.org

Widén-Nilsson, E., Halldin, S., and Xu, C. (2007) Global water balance modelling with WASMOD-M: Parameter estimation and regionalisation, Journal of Hydrology, 340, 105–118

Widén-Nilsson, E., Gong, L., Halldin, S., Xu, C. (2009) Model performance and parameter behaviour for varying time aggregations and evaluation criteria in the WASMOD-M global water balance model, *Water Resources Research*, 45(5), W05418

Winsemius, H.C., Schaefli, B., Montanari, A., Savenije, H.H.G. (2009) On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information, *Water Resources Research*, 45(12), W12422

Wisser, B., Fekete, M., Vörösmarty, C. J., Schumann, A. H. (2010) Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H), *Hydrology and Earth System Sciences*, 14, 1–24

Yamazaki, D. et al. (2017) A high-accuracy map of global terrain elevations, *Geophysical Research Letters*, 44, 5844-5853

Yang, D., Musiake, K. (2003) A continental scale hydrological model using the distributed approach and its application to Asia, *Hydrological Processes,* 17, 2855–2869

Yadav, M., Wagener, T., Gupta, H. (2007) Regionalisation of constraints on expected watershed response behaviour for improved predictions in ungauged basins, *Advances in Water Resources*, 30, 1756-1774

Yilmaz, K.K., Gupta, H.V., Wagener, T. (2008) A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrologic model, *Water Resources Research*, 44, W09417

Zhang, X., Srinivasan, R., Van Liew, M. (2008) Multi0site calibration of the SWAT model for hydrologic modelling, Transactions of the American Society of Agricultural and Biological Engineers, 51(6), 2039-2049

Zwarts, L. (2010) Will the Inner Niger Delta shrivel up due to climate change and water use upstream? A&W rapport 1537 Altenburg & Wymenga ecologisch onderzoek, Feanwâlden

Zwarts, L., Cissé, N., Diallo, M. (2005) *Hydrology of the Upper Niger*. In: Zwarts L, Beukering van P, Kone B, Wymenga E, editors. The Niger, a lifeline. Lelystad: RIZA/Wetlands International/IVM/A&W. p. 15-40

Zwarts, L. Grigoras, I. (2005) Flooding of the Inner Niger Delta. In: L. Zwarts, et al., eds. *The Niger, a lifeline: effective water management in the Upper Niger Basin*. RIZA: Wetlands International, 43–77